Correlation induced electrostatic effects in biomolecular systems

Dissertation

zur Erlangung des Grades des
Doktors der Naturwissenschaften
der Naturwissenschaftlich-Technischen Fakultäten
der Universität des Saarlandes

von
Sophie Weggler

Saarbrücken
Dezember 2010

Tag des Kolloquiums	27.04.2011
Dekan	Professor Dr. Holger Hermanns
Mitglieder des Prüfungsausschusses	Professor Dr. Volkard Helms (Vorsitzender)
	Professor Dr. Andreas Hildebrandt
	Professor Dr. Hans-Peter Lenhof
	Dr. Oliver Müller

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form in einem Verfahren zur Erlangung eines akademischen Grades vorgelegt.

Saarbrücken, 22. Dezember 2010

(Sophie Weggler)

Abstract
An understanding of electrostatic interactions in biomolecular systems is crucial for many applications in molecular biology. This thesis focuses on the theoretical modeling of two effects: first, the change in the dielectric properties of water due to hydrogen bond formation and second, the reentrant condensation of proteins induced by protein-metal ion complexation.

A nonlocal response theory is necessary to describe the dielectric effects of hydrogen bond formation. Correctly formulating this theory for a solvated biomolecule is challenging, because the biomolecule's cavity poses an obstacle for the water network. We develop a theory explicitly incorporating boundary conditions to describe the water network on the molecular surface. We implement an accurate and efficient finite difference solver, which offers the possibility to easily investigate different physically motivated boundary effects.

A detailed analysis of different nonlocal models reveals that, for the macroscopic behavior, the boundary conditions are of minor importance, while for a detailed understanding of the electrostatics near the molecular surface the correct modeling of the hydrogen bond formation is crucial.

Recent experimental findings describe a reentrant condensation of proteins in solutions of varying metal ion concentration. We present a heuristic model to account for the metal ion binding on the molecular surface which qualitatively and quantitatively explains the phase diagram of this condensation effect.

Kurze Zusammenfassung
In der vorliegenden Arbeit konzentrieren wir uns auf die Beschreibung elektrostatischer Phänomene in biomolekularen Systemen.

Zuerst untersuchen wir den Einfluss von Wasserstoffbrückenbindungen auf die dielektrischen Eigenschaften von Wasser. Dafür ist die Einführung eines nichtlokalen dielektrischen Operators notwendig. Die nichtlokale Reaktion des Wassers wird durch das gelöste Protein und der damit entstandenen Kavität maßgeblich beeinflusst. Wir entwickeln ein Differentialgleichungssystem, welches Veränderungen der dielektrischen Eigenschaften an der Molekülоberfläche explizit berücksichtigt. Um diese Randeffekte genauer zu analysieren und um unsere Modellgleichungen auf ionische Lösungen zu erweitern, implementieren wir ein modifiziertes Finite Differenzen Verfahren, welches sich, neben Effizienz, durch hohe Genauigkeit auszeichnet. Mit diesem Lösungsverfahren untersuchen wir erstmals verschiedene Wassermodelle. Die Analyse zeigt, dass die Veränderungen der Randbedingung an der Molekülоberfläche auf makroskopische Größen von untergeordneter Bedeutung sind, jedoch einen signifikanten Einfluss auf das elektrostatische Potential in der Nähe des Moleküls hat.

Des Weiteren betrachten wir einen kürzlich entdeckten Effekt in Proteinlösungen: die Bindungsaffinität von gelösten Metallionen induziert die Bildung von Protein-Metallionen-Komplexen. Diese können in Abhängigkeit der gelösten Ionenkonzentration kondensieren und wieder in Lösung gehen. In Analogie zu Protonierungsmodellen entwickeln wir eine Theorie zur Beschreibung der Komplexbildung. Erste Vergleiche mit Experimenten zeigen, dass das vorgeschlagene Modell den Kondensationseffekt qualitativ und quantitativ erklären kann.

Zusammenfasung

Die Funktionalität von Proteinen ist grundlegend für viele Reaktionen in lebenden Organismen, denn sie können andere Moleküle an sich binden und damit eine Wechselwirkung auslösen. Dieses sogenannte Schlüssel-Schloss-Prinzip wird in der medizinischen Forschung zur Wirkstoffentwicklung genutzt: hier sucht man unter anderem nach kleinen Molekülen, sogenannten Liganden, die sich an ein Protein anlagern um dessen fehlerhafte oder unerwünschte Reaktion im Organismus zu unterbinden.

Um Wirkstoffe zu finden oder zu kreieren, bedarf es eines hohen zeitlichen und finanziellen Aufwands. Mit Hilfe von theoretischen Modellen versucht man deshalb möglichst viele Erkenntnisse über potentielle Liganden zu erwerben. Hierbei ist es wichtig, die für die Wechselwirkung relevanten physikalischen Größen zu spezifizieren und sie so gut wie möglich in einem Modell abzubilden. Wegen ihrer Langreichweitigkeit und Stärke ist die Modellierung elektrostatischer Wechselwirkungen hierbei von größter Wichtigkeit. Dazu gehört nicht nur die Elektrostatik des Moleküls, sondern insbesondere auch der elektrostatische Einfluss seines Lösungsmittels. Diesen Einfluss bezeichnen wir im Folgenden als *dielektrische Antwort*. Da Wasser das natürliche Lösungsmittel von Molekülen in lebenden Organismen ist, ist dessen genaue Modellierung besonders wichtig.

Der erste Forschungsschwerpunkt der vorliegenden Arbeit ist die Modellierung der dielektrischen Eigenschaften von Wasser in biomolekularen Systemen. Wasser ist ein vielseitiges Lösungsmittel, denn neben seinen permanenten Dipolmomenten besitzen Wassermoleküle die Eigenschaft, Wasserstoffbrücken zu bilden, wodurch sie ein dreidimensionales Netzwerk ausbilden. Wegen dieses Netzwerks hängt die Bewegung eines einzelnen Moleküls von den Bewegungen der benachbarten Wassermoleküle ab. Anschaulich ist die Reaktion der Wassermoleküle also nicht nur bestimmt durch das elektrostatische Feld an dem Ort des Moleküls, sondern auch durch das in der näheren Umgebung herrschende Feld. Wir sprechen von einer *nichtlokalen* dielektrischen Antwort und damit von der *nichtlokalen* Elektrostatik. Sie steht im Gegensatz zu der klassischen, *lokalen* Elektrostatik, bei der die dielektrische Antwort allein von dem betrachteten Ort abhängt und wie folgt formuliert wird:

$$\boldsymbol{E}_{vac}(\boldsymbol{r}) = \varepsilon(\boldsymbol{r})\boldsymbol{E}_{wasser}(\boldsymbol{r})\,,$$

wobei E_{vac} das elektrostatische Feld im Vakuum, ε die dielektrische Funktion, und E_{wasser} das reale, durch die lokale dielektrische Antwort resultierende Feld bezeichnet.

In den 70er Jahren erweiterten russische Physiker obige Gleichung, um die dielektrischen Effekte des Wasserstoffbrückennetzwerks zu beschreiben. Die Erweiterung bestand in der Einführung eines dielektrischen Operators $\varepsilon(\boldsymbol{r}-\boldsymbol{r}')$, der die Reaktion an einer Position $\boldsymbol{r}$ in Wasser in Abhängigkeit der Position $\boldsymbol{r}'$ beschreibt:

$$\boldsymbol{E}_{vac}(\boldsymbol{r}) = \int_V \mathrm{d}\boldsymbol{r}'\varepsilon(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{E}_{wasser}(\boldsymbol{r}')\,,$$

dabei bezeichnet V den Raum, den das Wasser einnimmt. Wegen des Zusammenspiels dieses Integrals und der differentiellen Materialgleichungen der Elektrostatik konnte man die nichtlokalen Gleichungen nur für ausgewählte Modellsysteme lösen. A. Hildebrandt et al. ermöglichten 2007 mit der Einführung einer rein differentiellen Schreibweise die nichtlokale elektrostatische Theorie auf kompliziertere Modellsysteme, wie zum Beispiel in Wasser gelöste Proteine, anzuwenden.

Basierend auf dieser Umformulierung des integro-differentiellen Systems wurden zwei unterschiedliche Wassermodelle entwickelt. Dies ist der Ausgangspunkt unserer Untersuchungen.

Obwohl beide Wassermodelle auf dem gleichen nichtlokalen Operator basieren, sind die resultie-

renden Differentialgleichungen verschieden. Erstes Ziel der vorliegenden Arbeit ist es, ein besseres Verständnis der unterschiedlichen Gleichungssysteme zu erlangen. Ausgehend von der mathematischen Idee die integrale Gleichung als Lösung eines Differentialgleichungssystems zu verstehen, untersuchen wir erstmals die physikalische Bedeutung dieser Umformulierung: wir führen ein neues physikalisch interpretierbares Feld, das sogenannte Korrelationsfeld, ein. Dieses Feld trägt der durch die Wasserstoffbrückenbindungen hervorgerufenen Nichtlokalität Rechnung. Stellen wir uns das Wassernetzwerk in der Nähe des gelösten Moleküls vor, dann wird unmittelbar klar, dass die Oberfläche des Moleküls ein Hindernis für das Wassernetzwerk darstellt. Eine Umorientierung der Wasserstoffbrücken an der Molekülоberfläche verändert das Korrelationsfeld. Mit der expliziten Einführung dieses Feldes können wir nun die Unterschiede der beiden Wassermodelle physikalisch interpretieren: das eine Modell legt eine kontinuierliche Fortsetzung des Korrelationsfeldes zugrunde, wohingegen das andere ein verschwindendes Korrelationsfeld an der Molekülоberfläche fordert.

Diese Überlegungen führen zu der Frage, ob es eine Möglichkeit gibt, explizit das Verhalten des Korrelationsfeldes an der Oberfläche des Moleküls vorzugeben. Solche Randwerte können zum Beispiel theoretisch motiviert oder aus einem Experiment abgeschätzt worden sein. In der vorliegenden Arbeit präsentieren wir eine neue differentielle Formulierung der nichtlokalen Theorie, in der die Vorgabe von sogenannten Randbedingungen das Verhalten des Korrelationsfeldes an der Oberfläche fixieren.

Um die Effekte verschiedener Randbedingungen des Korrelationsfeldes effizient zu untersuchen, ist ein numerisches Verfahren notwendig, das schnell umzusetzen und einfach erweiterbar ist. Ergänzend zu den theoretischen Überlegungen entwickeln wir in Kooperation mit V. Rutka eine modifizierte Finite Differenzen Methode (FDM). Unsere FDM nutzt die sogenannte Explizit Jump Immersed Interface Methode (EJIIM), die sich neben Effizienz durch hohe Genauigkeit der Lösungen auszeichnet. Bei diesem Verfahren werden Sprungbedingungen in den Gleichungen durch Korrekturterme direkt modelliert. Durch die Veränderung der dielektrischen Antwort an der Molekülоberfläche treten definierte Sprünge in den elektrostatischen Gleichungen auf, was EJIIM zu einer geeigneten, leistungsfähigen numerischen Methode macht.

Für die Anwendung der EJIIM sind zwei Vorarbeiten notwendig. Zum einen, bedarf es einer exakten Beschreibung der Molekülоberfläche in einem dreidimensionalen, kartesischen Gitter. In dieser Arbeit stellen wir einen neu entwickelten Algorithmus vor, der eine Gitter-basierte Molekülоberfläche berechnet. Zum anderen erfordert das FDM eine Abschätzung der gesuchten Felder an den Rändern des benutzten Gitters. Wir leiten eine exzellente Näherung her. Eine Analyse dieser Randwertnäherung zeigt, dass sie in dem für uns interessanten Parameterbereich nicht nur am Rand, sondern im gesamten Lösungsmittelgebiet gilt. Dies eröffnet neue Anwendungsmöglichkeiten der nichtlokalen Elektrostatik, wie zum Beispiel die zeitabhängige Simulation der Bewegung eines Liganden in der Nähe der Proteinbindetasche.

Mit der EJIIM studieren wir unterschiedliche Wassermodelle. Die Untersuchung ergibt, dass die nichtlokale Elektrostatik im Vergleich zu der lokalen Modellierung ein größeres, elektrostatisches Potential im Außenraum und ein kleineres Reaktionsfeldpotential im Molekülinneren aufweist. Die Größenordnung ist bei allen nichtlokalen Modellen gleich. Unterschiede ergeben sich lokal bei den Potentialen in direkter Nähe zur Proteinoberfläche.

Der zweite Forschungsschwerpunkt dieser Arbeit ist ein spezieller Kondensationseffekt, der kürzlich bei Protein-Metallionen-Lösungen gefunden wurde: die sogenannte reentrante Kondensation. Reentrante Kondensation bezeichnet den Effekt, dass sich Proteine negativer Gesamtladung bei Zugabe einer kritischen Konzentration c^* hochgeladener Metallionen zusammenlagern. Bei Präsenz einer zweiten, kritischen Metallionenkonzentration $c^{**} > c^*$ geht das Kondensat wieder in Lösung. Eine genauere Analyse von Metallionen-Protein Wechselwirkungen legt den Schluss nahe, dass es sich hierbei um eine spezifische Metallionen-Protein-Komplexierung handelt. Diese versuchen wir

durch ein einfaches Modell darzustellen: wir behandeln die Metallionen als Liganden, die sich an vordefinierten Bindungsstellen auf der Proteinoberfläche anlagern. In Anlehnung an das Vorgehen bei Protonierungsimulationen führen wir Affinitätskonstanten ein, die das Anlagern in Abhängigkeit der Ionenkonzentration energetisch begünstigen. Mit der Postulation, dass, erst wenn die effektive Gesamtladung der Protein-Metallionen-Komplexe verschwindet, die Kondensation durch attraktive Wechselwirkungen einsetzt, kann diese Modellvorstellung den Kondensationseffekt konsistent erklären.

Die theoretischen Überlegungen setzen wir numerisch um: wir implementieren ein modifiziertes Titrationsverfahren, welches mit Hilfe einer Monte Carlo Simulation den Zustandsraum abtastet und den energetisch günstigsten Besetzungszustand der Metallionen auf der Proteinoberfläche findet. Mit Hilfe dieses Modells berechnen wir die effektive Gesamtladung eines Protein-Metallionen-Komplexes. Der qualitative und quantitative Vergleich der ersten Phasengrenze und der effektiven Ladungskurve, sowie der qualitative Vergleich mit Zeta-Potential-Kurven, zeigt sehr gute Übereinstimmung.

Insgesamt trägt diese Arbeit zu einem größeren Verständnis von Korrelationseffekten elektrostatischer Phänomene, die in biomolekularen Systemen auftreten, bei. Die vorgeschlagenen Modelle und theoretischen Betrachtungen eröffnen interessante Fragestellungen für zukünftige, innovative Forschungsarbeiten. Des Weiteren bilden die von uns entwickelten numerischen Methoden die Basis, theoretische Überlegungen effizient umzusetzen und zu untersuchen.

Zum Glück

... hatte ich während der letzten vier Jahre an vielen Ecken und Enden Hilfe und dafür möchte ich mich aus tiefsten Herzen bedanken. Da den Leser noch rund 200 Seiten erwarten, muss ich mich hier kürzer fassen als mir lieb ist und es bei der reinen Erwähnung von Namen oder einer kleinen und damit in keinster Weise vollständigen, Erklärung belassen.

Meinem Chef, Andreas Hildebrandt gilt zu allererst mein Dank, denn er hat mich mit einem sehr interessanten und vielseitigen Thema betraut, an dem ich ungemein viel gelernt habe. Er hat mich in meiner Arbeit und dem Vertrauen in mein Können immer bestärkt, und wenn ich mal den Wald vor lauter Bäumen nicht sehen konnte, hatte er die richtigen Worte und Ideen parat.

Ohne meine Koorperation mit Vita Rutka wäre ein Teil dieser Arbeit niemals zustande gekommen. Diese Kooperation war nicht nur erfolgreich und hat mich nicht nur in neue, numerische Methoden eingeführt, sie hat mir darüber hinaus sehr viel Spass gemacht.

Sergej Rjasanow, Manfred Lücke, Hans-Peter Lenhof und Frank Döbrich möchte ich meinen Dank aussprechen, denn durch ihre Unterstützung konnte ich mein letztes Jahr über ein Stipendium der Universität des Saarlandes finanzieren.

Eine nicht zu unterschätzende Hilfe in der gesamten Zeit der Promotion waren meine Kollegen von den Arbeitsgruppen Lenhof und Hildebrandt, die mich in den kleinen Kreis der Bioinformatik aufgenommen haben und mir immer geduldig bei meinen Programmierversuchen geholfen haben. Besonderer Dank gilt dabei meinem “Mitbewohner” René Hussong. Unsere gemeinsame Zeit und die Spaziergänge zur Sportlermensa werde ich in Erinnerung behalten.

Das Ende der Promotion ist das Ende meiner Studienzeit in Saarbrücken. In dieser gesamten Zeit ist meine Familie für mich da gewesen: nicht nur finanziell sondern in allen Lebenslagen war ich ihrer Unterstützung sicher. Das hat mir meinen Lebensweg erleichtert und mich bereichert. Ein ganz besonderer Dank gilt hier meiner Schwester Lucy, die sich schon allein als unermüdliche Korrekturleserin meiner Arbeit einen Ehrentitel verdient hat. Bezüglich Korrekturen will ich an dieser Stelle auch Daniel Stöckel danken. Ich hätte keine besseren “Kritiker” haben können.

Darüber hinaus möchte ich allen nachfolgenden Personen danken: Caroline Fasel, Michael Ziller, Fajun Zhang und Frank Schreiber, Oliver Kohlbacher, Dirk Neumann, Holger Kohr und Martin Riplinger, Veronika Mueller, Richards Grzibovskis, Jan Baumbach, Christina Backes und Benny Kneissl, meinen “alten” Studienkollegen, die mich auch während der Promotion begleitet haben, und Philipp Schäfer.

Contents

1 Introduction **1**

2 Physical background **9**
2.1 The Maxwell equations in vacuum . . . 9
2.1.1 The Laplace operator . . . 12
2.2 The Maxwell equations in dense media . . . 14
2.2.1 Monopole and dipole moment of a charge distribution . . . 14
2.2.2 The macroscopic material equations . . . 16
2.2.3 The dielectric operator . . . 18
2.3 Coulomb interaction energy and field energy . . . 20

3 The biomolecular system **23**
3.1 Model for biomolecules . . . 23
3.1.1 Charge distribution of biomolecules . . . 24
3.1.2 Surface description of biomolecules . . . 25
3.1.3 Dielectric response of proteins . . . 26
3.1.4 Solvation energy . . . 27
3.2 Water - the main solvent of biomolecules . . . 30
3.2.1 Characteristics of water . . . 30
3.2.2 The dielectric response of water . . . 31
3.2.3 The Lorentzian model of water . . . 36
3.3 The Lorentzian model for biomolecules in water . . . 38
3.3.1 The dielectric function in the biomolecular system . . . 39
3.3.2 The Maxwell equations of the biomolecular system . . . 41
3.3.3 Application to spherically symmetric systems . . . 44

4 Reformulation of the integro-differential equation **51**
4.1 The Yukawa operator . . . 52
4.2 Boundary and transmission problem of the Yukawa operator . . . 53
4.3 Generalization of the correlation field . . . 56
4.3.1 Newton potential approach . . . 58
4.3.2 Dirichlet problem . . . 61
4.4 The correlation field in differential form . . . 64
4.5 Conclusion . . . 65

5 Scalar approximations of the Lorentzian model **67**
5.1 Decomposition of the Dirichlet Vector Model . . . 68
5.1.1 Screening length of orthogonal solutions . . . 71
5.1.2 Existence of orthogonal solutions . . . 71
5.1.3 Dirichlet Scalar Model . . . 73
5.2 Decomposition of the Newton Vector Model . . . 74
5.2.1 Characterization of $\mathbf{F}_{\mathrm{N}}^{hom}$ . . . 74

5.2.2 Convolutions as "smoothing" operations . . . 75
5.2.3 Newton Scalar Model . . . 77
5.3 Discussion and comparison . . . 78

6 Numerics **83**
6.1 BEM . . . 85
6.1.1 Representation formulas . . . 86
6.1.2 Boundary integral equations . . . 86
6.1.3 Surface triangulation . . . 88
6.2 EJIIM . . . 88
6.2.1 EJIIM discretization . . . 88
6.2.2 Approximating the jump variables . . . 91
6.2.3 Solving the discrete system . . . 93
6.2.4 Boundary condition . . . 94
6.3 Surface generation . . . 99
6.3.1 Definitions . . . 99
6.3.2 Construction of the grid-based SAS . . . 103
6.3.3 Construction of the grid-based SES . . . 108
6.3.4 Final remarks . . . 111
6.3.5 The marching cubes algorithm . . . 113
6.4 Numerical comparison of EJIIM and BEM . . . 115
6.4.1 Convergence towards analytical solution . . . 115
6.4.2 Comparison on molecular geometries . . . 118
6.5 Conclusion . . . 123

7 Model comparison and applications **125**
7.1 Comparison on the test set . . . 126
7.1.1 Solvation energy . . . 126
7.1.2 Electrostatic potential . . . 127
7.2 Nonlocal electrostatic application to biomolecules . . . 131
7.2.1 Trypsin . . . 132
7.2.2 SIV/HIV capsid - primate TRIM5α interaction . . . 136
7.2.3 Ovalbumin . . . 138
7.3 Summary . . . 139

8 Electrostatics of biomolecules in ionic solution **141**
8.1 Mean field theory of ionic solutions . . . 141
8.2 Electrostatics in ionic solutions . . . 143
8.2.1 Dirichlet Scalar Model with ionic screening . . . 144
8.2.2 Conclusion . . . 148
8.3 Beyond mean field theory . . . 149
8.3.1 Experimental findings . . . 150
8.3.2 Theoretical model . . . 153
8.3.3 Implementation . . . 157
8.3.4 Results . . . 159
8.3.5 Discussion . . . 160
8.3.6 Conclusion . . . 164

9 Summary and outlook **167**

10 Appendix **173**
10.1 Notations relating to spaces of integrable functions 173
10.1.1 Sobolev spaces for open subsets Ω of $\mathbb{R}^d$ 173
10.1.2 Sobolev spaces on boundaries Γ 175
10.1.3 Spaces of the Helmholtz decomposition 176
10.2 Continuity condition of the electrostatic field and its potential 177
10.3 Titration of amino acids 178
10.3.1 Dissociation constant and the pK_a value of an amino acid 178
10.3.2 pK_a values in proteins 180
10.3.3 Free energy difference for given protonation state 182
10.3.4 Monte Carlo simulation of protein titration 182
10.4 Three possibilities to calculate binding energies 184

List of Figures

1.1 Steps in drug development . . . 1
1.2 Protein in water I . . . 2
1.3 Water dipole . . . 2
1.4 Hydrogen bonds . . . 3
1.5 Water network near the molecular surface . . . 4
1.6 Protein in water II . . . 5
1.7 Reentrant condensation . . . 6

2.1 Maxwell equations . . . 10
2.2 Transmission condition I . . . 11
2.3 Transmission condition II . . . 12
2.4 Taylor series I . . . 15
2.5 Electric dipole . . . 15
2.6 Taylor series II . . . 16

3.1 Proteins are chains of amino acids . . . 24
3.2 Domains of the molecular system . . . 25
3.3 Molecular surface descriptions . . . 26
3.4 Molecular and reaction field potential . . . 27
3.5 Water network near the molecular surface . . . 29
3.6 Water molecule . . . 30
3.7 Correlation length of the water network . . . 31
3.8 Possible shapes of $\hat{\varepsilon}(k)$ . . . 35
3.9 Nonlocal response on the dielectric boundary . . . 40
3.10 Nonlocal electrostatics for spherical symmetry . . . 44
3.11 Electrostatic potential for spherical symmetry . . . 49
3.12 The free energy of solvation of monoatomic ions . . . 50

4.1 Overview on models for nonlocal electrostatics I . . . 57

5.1 Overview on models for nonlocal electrostatics II . . . 67
5.2 Helmholtz decomposition . . . 68
5.3 The biomolecular surface in spherical coordinates . . . 72
5.4 Smoothing operation I . . . 76
5.5 Smoothing operation II . . . 76
5.6 Characteristic constant B of the Born model . . . 80

6.1 Overview on numerical models for nonlocal electrostatics . . . 84
6.2 Regular and irregular points . . . 89
6.3 Irregular point with two intersections . . . 90
6.4 Stencil definition . . . 92
6.5 Boundary value approximation . . . 94
6.6 Surface expansion . . . 95

6.7 Restrictions for discrete surface description . . . 100
6.8 Grid based representation of an atom sphere . . . 102
6.9 Atom connectivity . . . 103
6.10 Merging scenarios I . . . 107
6.11 SAS generation . . . 107
6.12 Sketch of the SES generation . . . 108
6.13 De Morgan's law . . . 109
6.14 Probe sphere versus atom sphere . . . 109
6.15 Merging scenarios II . . . 111
6.16 SES construction . . . 112
6.17 Marching cubes configurations . . . 113
6.18 Triangulation I . . . 114
6.19 Triangulation II . . . 114
6.20 Convergence of EJIIM and BEM I . . . 116
6.21 Convergence of EJIIM and BEM II . . . 116
6.22 Convergence of EJIIM and BEM III . . . 118
6.23 Contour lines of ϕ_Σ for FUDPOJ and GUANCH01 . . . 122

7.1 Convergence of EJIIM and BEM IV . . . 125
7.2 Contour lines of ϕ_{mol} . . . 128
7.3 Contour lines of ϕ_Ω and ϕ_Ω^{reac} of BRMW1 . . . 128
7.4 Contour lines of ϕ_Σ of BRMW1 . . . 129
7.5 Contour lines of ϕ of FUDPOJ . . . 130
7.6 Protein surfaces of trypsin and ovalbumin . . . 132
7.7 Approximated boundary condition . . . 133
7.8 Contour lines of ϕ_{mol} of trypsin . . . 134
7.9 Contour lines of ϕ of trypsin . . . 134
7.10 Isosurfaces of trypsin . . . 136
7.11 CA variants of HIV . . . 137
7.12 Electrostatic potential of the CA variants of HIV . . . 138
7.13 Isosurfaces of ovalbumin . . . 139

8.1 Protein in water II . . . 142
8.2 Born solution in the DSM_LPB . . . 146
8.3 Application of the DSM_LPB on small molecules . . . 148
8.4 Theoretical phase diagram . . . 149
8.5 Structural difference of protein and DNA . . . 150
8.6 Experimental phase diagram . . . 151
8.7 Thermodynamic cycle . . . 156
8.8 Binding curves of BSA . . . 160
8.9 Effective charge of BSA . . . 161
8.10 Determination of fit parameter α . . . 162
8.11 Comparison of zeta potential and effective charge of BSA . . . 164

10.1 Continuity of the electrostatic potential . . . 177
10.2 A sequence of amino acids . . . 178
10.3 Arginine . . . 179
10.4 Free energy cycle . . . 181
10.5 Titration: lysozyme . . . 183
10.6 Binding energy I . . . 184

10.7 Two different solvation scenarios 185
10.8 Binding energy II 185
10.9 Binding energy III 186
10.10Binding energy IV 187

List of Tables

5.1 Born models for different nonlocal models 79
5.2 Solvation free energy for the Born models 80

6.1 Notations 101
6.2 Edge label 102
6.3 Molecule edge types and addition operator $\oplus_A$ 106
6.4 Addition operator $\oplus_P$ 111
6.5 Numerical data of the test set 118
6.6 Numerical convergence I 119
6.7 Numerical convergence II 120
6.8 Approximation of the exterior boundary condition 121
6.9 Computational costs for AN05A 123

7.1 Solvation energies for the test set 127

8.1 Experimental results of the phase transitions 152
8.2 Nomenclature used in titration theory 153
8.3 Input quantities for the electrostatic calculations 159
8.4 Experimental and numerical results 163

10.1 Standard pK_a values 180

Chapter 1

Introduction

Proteins are the main actors in living organisms and their interactions are fundamental to many processes in molecular biology. A characteristic of proteins that allows a diverse set of functions is their ability to bind molecules, such as other proteins, DNA, or ligands.

In particular, protein-ligand docking is of great interest in pharmaceutical studies and drug development: the idea is to search for small ligands that - bound to the disease-related protein - suppress the functionality of the latter. Searching for new, medical therapies, enabled by suitable ligands, is an important, but, very expensive, field of research of our time. Fig. 1.1 depicts the various steps in drug development with respect to temporal and financial effort.

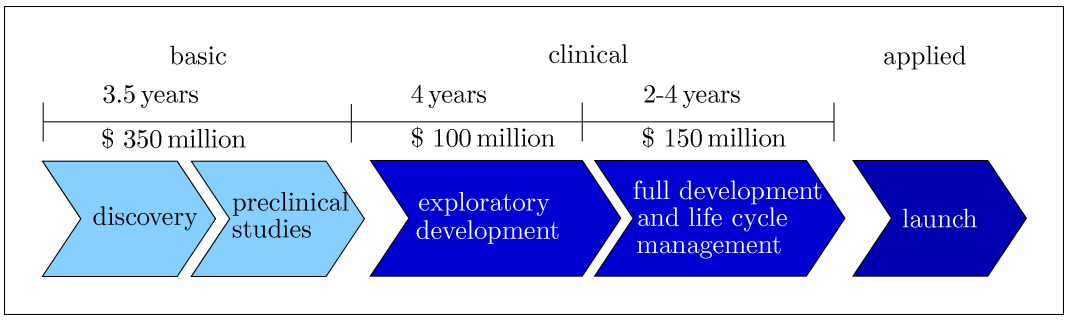


Fig. 1.1: Steps in drug development with their temporal and financial effort [89].

In order to reduce the costs as well as the time needed for testing novel drug candidates in the clinical phase, it is of great importance to carefully study the required features of a potential drug. In the preclinical phase, research is focused on the rational computer-aided drug design with the aim to find the most promising candidates for a further study. A crucial step towards an effective drug is to accurately predict biochemical interactions and therefore, effort is made to understand the basics of molecular interactions from the theoretical side.

Electrostatic interactions play a particularly important role in searching for suitable docking partners, i.e., drug candidates, because of their long-range character and the strong attraction, which is present between two contrary charged regions. For correctly understanding the electrostatic interactions of biomolecules it is, however, essential to take into account its native solvent, namely water. The permanent dipole moment of the water molecules crucially changes the electrostatics of the immersed molecules. Further, the ability of water to form a hydrogen bond network causes this solvent to be rather dynamic and active than inert in the presence of a protein. Water is, for instance, known to electrostatically stabilize the native structure of biomolecules and to contribute significantly to protein folding [36, 152]. It also plays a crucial role in the functionality of proteins, see [20] and citations given therein.

Understanding the complex interplay between water and proteins as well as the functionality of the latter implies a deeper knowledge on the electrostatic interactions apparent in the solvent.

The standard theory, which has been found to describe the electrostatic fields caused by all the partial charges present in a system, is given by the Maxwell equations. As illustrated in Fig. 1.2, the

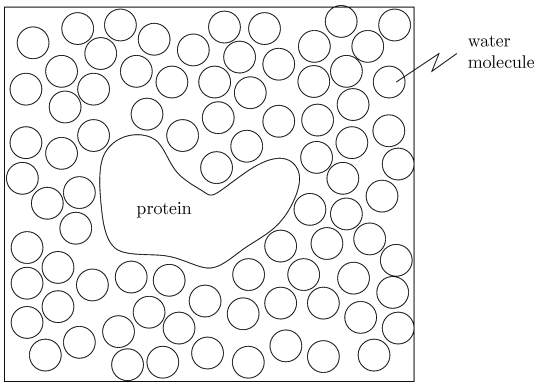


Fig. 1.2: In the native state, the protein is immersed in water.

protein is naturally surrounded by a multitude of water molecules. The protein as well as the mobile water molecules carry partial charges, which cause them to interact with each other in a highly dynamic and flexible way to reach an equilibrium state. A prediction of these interactions, with simultaneously considering all particles in the system, would mean to apply the time dependent Maxwell equations on a many-particle problem of the order $\sim 10^{23}$ particles/liter. Such an explicit representation of each water molecule is often not feasible and further, is often not necessary: because of their high density the water molecules can be assumed as a continuous medium on the macroscopic length scale, in particular when considering a time average.

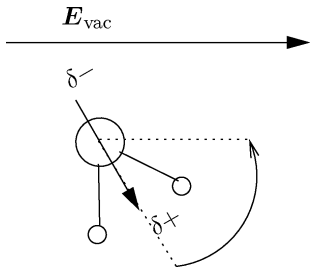


Fig. 1.3: A water dipole rotates to minimize its potential energy in the electrostatic field.

The ability to form hydrogen bonds is often not explicitly incorporated in a theoretical model, but the water molecules are treated as small, rotatable dipoles in the electrostatic field, which originates from the protein's charges in vacuum. The reorientation of a single water molecule in the electrostatic field $\boldsymbol{E}_{\mathrm{vac}}$ is depicted in Fig. 1.3.

The macroscopic effect of all the polar solvent molecules, also known as *dielectric response*, results in a weakening of the electrostatic field compared to its magnitude in vacuum [61]. In fact, most biomolecular studies assume a simple continuum model of water, where a constant factor of $\varepsilon_{\mathrm{wat}}/\varepsilon_{\mathrm{vac}} = 78.5$ is used to describe the dielectric effect. This means that the electrostatic field $\boldsymbol{E}_{\mathrm{wat}}$ of the protein in water is weakened by a factor of 78.5 compared to the electrostatic field $\boldsymbol{E}_{\mathrm{vac}}$ of the protein in vacuum,

$$\boldsymbol{E}_{\mathrm{vac}} = \frac{\varepsilon_{\mathrm{wat}}}{\varepsilon_{\mathrm{vac}}} \boldsymbol{E}_{\mathrm{wat}} \,.$$

Whether this assumption is adequate, highly depends on the aim of the biomolecular study.

As we have discussed before, the dipolar character of water *and* the hydrogen bonds between the water molecules are not negligible when the functionality of proteins in their native state is the focus of research. Fig. 1.4 shows that a water molecule tries to correlate with four neighboring water molecules mediated by hydrogen bonds. This correlation influences the undisturbed rotation of each water molecule in the electrostatic field.

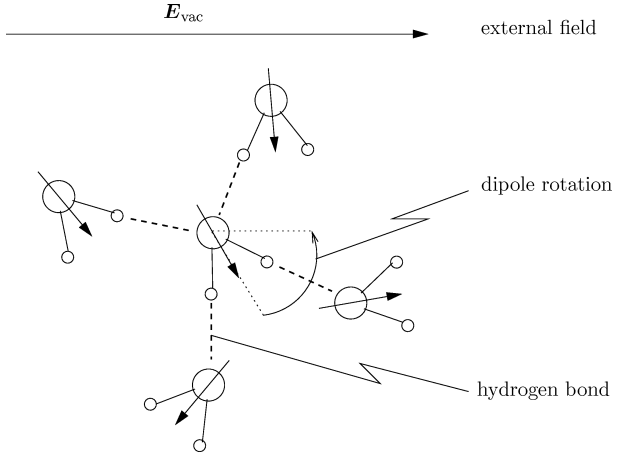


Fig. 1.4: The water molecule correlates via hydrogen bonds with its next neighbors.

First attempts to describe the structural effects of the hydrogen bond network within the electrostatic continuum theory go back at least forty years, when the simple continuum theory of constant dielectric response was extended to an integral form:

$$\boldsymbol{E}_{\mathrm{vac}}(\boldsymbol{r}) = \int \mathrm{d}\boldsymbol{r}' \varepsilon(\boldsymbol{r} - \boldsymbol{r}') \boldsymbol{E}_{\mathrm{wat}}(\boldsymbol{r}')$$

The integral representation of the dielectric response captures the spatial dependencies of the solvent's reaction to an external field [35]. These spatial dependencies originate from the coordinative hydrogen bonds of every water molecule with the neighboring molecules and therefore it is a *nonlocal response*.

Although the concepts to describe water in the nonlocal response theory has been established since then, the theory could only be applied to simple, symmetric systems because of the interplay of integral and differential equations. After A. Hildebrandt and coworkers achieved a reformulation of the integro-differential system to a system of complicated, however purely differential, equations, it is now possible to solve for the nonlocal electrostatics of non-trivial systems [52].

The first part of this work is based on the ideas given in [52, 53] and [39, 40] to describe the water correlations within the Lorentzian water model which is the simplest nonlocal model for $\varepsilon(\boldsymbol{r} - \boldsymbol{r}')$. The purely differential formulation of the electrostatic equations with a Lorentzian water response offers the possibility (a) to thoroughly investigate and interpret the physical meaning of the underlying nonlocal formulation, (b) to debate approximations for an efficient numerical solver and further (c) to discuss modifications and extensions in order to describe proteins in their native state:

In the following, we address these points in more detail, since all of them are tackled within this thesis.

(a) The original Lorentzian model of nonlocal response is actually a theory of the water in the bulk, i.e., there is *no* disturbing boundary and the whole space is filled with water. Thus, to

apply the nonlocal response theory to biomolecules immersed in water, we have to provide the tools for defining the behavior of water on the molecular surface. Here, the solid-liquid interface confines the water network: the spatial cavity caused by the immersed biomolecule as well as the change in the dielectric response when crossing the biomolecular surface naturally result in a reformation of the hydrogen bonds near the surface as it is shown in Fig. 1.5.

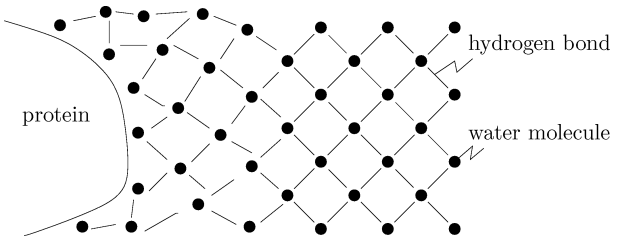


Fig. 1.5: The water network is disturbed when approaching the molecular surface.

In this work, we derive a novel formulation of the nonlocal water response, where in addition to the volume integral, which has been presented above, boundary integrals determine the dielectric response in the solvent and in particular on the surface. This gives us - at least theoretically - the possibility to impose a specific, variable behavior of the water correlations on the molecular surface. It offers, for instance, a consideration of "first shell" effects [51] or the differentiation of local variations of the hydrogen bond network due to polar and nonpolar regions.

With this general integral formulation we gain a deeper understanding of two former nonlocal approaches: the vectorial model that has been derived and numerically solved by C. Fasel [40], is based on the original Lorentzian water model, i.e., it ignores an explicit change in the water network when approaching the molecular surface. In contrast, the second model proposed by A. Hildebrandt, incorporates - in its vectorial formulation - a stiff water network on the molecular surface. Apart from the interpretation of the vectorial models, we find the physical meaning of the approximations made by A. Hildebrandt in [53] to obtain a purely potential formulation of the nonlocal electrostatic equations.

(b) Both, the vectorial and the scalar, models are numerically solved using the Boundary Element method. Although this method yields highly accurate results, it is difficult and sometimes impossible to adapt the method to modifications. For instance, considering Boltzmann-distributed salt ions in the solvent results in an extension from linear to nonlinear partial differential equations, which can be solved with the Boundary Element method only with a considerable additional effort. A further bottleneck of the Boundary Element method is the requirement of a smooth, but coarse, surface triangulation because of limiting memory capacities. Coping with these requirements is not a simple task, when considering the complex shape, which is built up by hundreds of atoms the biomolecule is composed of. This can already be seen from the fact that up to now a fully automatic Boundary Element solver with integrated mesh generator for biomolecular computations is not available [83]. However, such an automatic solver is necessary for an integration of the nonlocal electrostatic models into a biomolecular framework and thus to offer a broad applicability of the nonlocal theory.

In order to avoid the inflexibility and the sensitive input requirements of the Boundary Element method, the second focus of this work is to develop an alternative, numerical tool, which solves the nonlocal differential equations efficiently and stably, and most importantly, which is easily extensible.

Using finite difference methods, we can draw upon many years of experience for computing biomolecular electrostatics [8, 11, 58, 85]. This, in principle, supports the idea to develop a finite different scheme for nonlocal electrostatics. However, several essential features of the electrostatic field on the interface of the molecule and the solvent are left untreated in traditional finite difference solvers [83]. Therefore, we decided to develop a *modified* finite difference scheme, the Explicit Jump Immersed Interface method, which explicitly accounts for the electrostatics on the molecular surface [143, 144]. On the one hand, we can take advantage of the stability and the effortless extensibility as well as of well-established numerical tools for solving the algebraic equations. On the other hand, the Explicit Jump Immersed Interface method provides highly accurate results even on the molecular surface.

With the successful implementation of this robust finite difference solver and the stable, automatic generation of required input information, we provide a new tool to study the nonlocal electrostatic equations and pave the way for a broader application. As elaborated in (a), we aim at a deeper understanding of the Lorentzian water theory giving rise to modified models, which can now be easily solved and compared.

(c) As illustrated in Fig. 1.6, the physiological conditions of native proteins are not only given by water, but also by additionally solvated ions, which determine the natural salinity. The Poisson-Boltzmann theory correctly describes the electrostatic field of biomolecules in ionic solutions for mono- and divalent ions of low concentration, i.e., in the case of negligible ion correlations [56]. Intuitively, it is clear that the ions distribute in a way to minimize their potential energy and this in turn results in an effect similar to the rotation of the dipoles which we discussed above: the electrostatic field is weakened in comparison to its magnitude in vacuum. The effect that solvated ions decrease the magnitude of the electrostatic field is known as *ionic screening*.

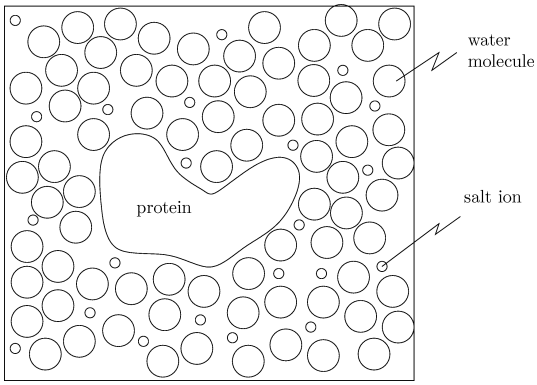


Fig. 1.6: In the native state, the protein is immersed in ionic water.

In order to further approach a realistic description of native proteins, we consider the effects of dilute, ionic solutions in addition to the nonlocal water response. This is accomplished by an incorporation of the linearized Poisson-Boltzmann theory into the nonlocal electrostatic equations. With the use of the finite difference method, such an extension is easily possible in the nonlocal framework.

The incorporation of salt ions into the theory of correlated water shows an interesting effect: the screening of the ions is a counterpart to the water network. Depending on the strength of both, the ionic screening and the water correlations, one of the two effects dominates and determines the overall behavior of the biomolecular system.

The description of the nonlocal water-water correlations as discussed before constitutes the main part of this work. But of course, not only water-water correlations can cause an extraordinary behavior in protein solutions. We complete this work with an investigation of specific protein-metal ion correlations treating the solvent as an unstructured medium, i.e., water is characterized by a constant dielectric response. Specific protein-metal ion interactions are important in protein solutions as they stabilize the protein's globular structure [10,49,100,131] and as they play a crucial role in the protein folding process. Specific protein-metal ion correlations can even cause protein aggregation and re-dissolution [150]. The latter phenomenon is called reentrant condensation and describes the effect that solvated protein precipitates due to adding a critical concentration c^* of trivalent metal ions. The solution once again re-dissolves, when adding a second critical concentration $c^{**} > c^*$ of trivalent metal ions. This can be seen on the change in the turbidity of the solution in Fig. 1.7. Such a controllable and reversible aggregation could be important for the drug development of human diseases, where the agglomeration of protein plays a crucial role, i.e., Alzheimer disease.

Fig. 1.7: Increasing the metal ion concentration (Y^{3+}) causes the protein solution (Bovine Serum albumin) to aggregate and to re-dissolve.

Specific protein-metal ion correlations are not treated in the Poisson-Boltzmann equation as within this theory the reaction of solvated ions is only captured by a mean-field approximation. We propose and develop a model, which accounts for these specific protein-metal ion correlations. The model is based on the introduction of binding affinities of the metal ions to surface exposed amino acids in analogy to titration simulations of the protonation state of proteins [111].

The outline of this work is as follows: in Chapter 2 we give a detailed introduction to the electrostatic Maxwell equations, as these are the basis not only for the conventional description of electrostatic phenomena in dielectrics, but they are also the basis for the nonlocal formulation which is deduced in Chapter 3. Within Chapter 3, we define the description of the biomolecule used in this work, i.e., its charge distribution, the molecular surface, and its dielectric response. Further, we introduce the details of water as the native solvent of biomolecules. We introduce the Lorentzian model to account for nonlocal water correlations and discuss analytically solvable systems.

Taking into account the idea to reduce the integro-differential system of nonlocal electrostatics to a system of purely differential equations, we develop a novel, general formulation of nonlocal electrostatics in Chapter 4. This formulation is the starting point for a detailed discussion on two different vectorial models. Chapter 5 is devoted to suitable approximations allowing for purely potential formulations of nonlocal electrostatics. For an extensive study of the nonlocal electrostatics of non-trivial geometries, we introduce two numerical tools in Chapter 6, the Boundary Element method and the Explicit Jump Immersed Interface method. As the latter has been developed for nonlocal electrostatics during this work, we discuss the basic idea and the numerical implementation in detail. This chapter further introduces the numerical tools to generate all the input requirements of the Explicit Jump Immersed Interface method: an estimation of the solution values on the boundary of the box used for the finite difference calculation as well as the discrete representation of the molecular surface in a 3-dimensional Cartesian grid. After a detailed discussion on the performance of the two numerical methods, we apply and compare the previously derived nonlocal models on a set of small molecules and two proteins in Chapter 7.

In Chapter 8 we analyze the mean-field effect of salt additionally immersed in correlated water. In order to demonstrate specific effects which are caused by correlations different from water-water correlation we finally focus on a recently discovered condensation effect in protein solutions. We demonstrate that this non-trivial effect is well described by a modified titration program, which we develop and apply in this chapter.

Chapter 2

Physical background

In this section, we present the so called *Maxwell equations*, which describe all dynamic and static phenomena that involve charges. Since the focus of this work is the electrostatic nature of molecules in solution, we concentrate on the time-independent formulation.

Section 2.1 introduces the Maxwell equations in vacuum, within which all charges of the system are treated explicitly. When considering a continuous, dense medium composed of a high number of molecules, an explicit treatment of all the involved charges is nearly impossible. In particular, in the context of solvated biomolecules we have to cope with this problem. Section 2.2 focuses on the so called *material equations*, where the effect of the charges is separated into two parts, an explicit and a mean-field part. A discussion on the electrostatic (field) energy in Section 2.3 closes the chapter on the physical background.

2.1 The Maxwell equations in vacuum

Let us consider an arbitrary, static charge density ρ in vacuum. Then, the electrostatic field $\boldsymbol{E}$ satisfies the following differential equations [61, 97]

$$\boldsymbol{\nabla} \cdot \boldsymbol{E}(\boldsymbol{r}) = \frac{1}{\varepsilon_0}\rho(\boldsymbol{r}), \qquad \boldsymbol{r} \in \mathbb{R}^3 \tag{2.1}$$

$$\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{0}, \qquad \boldsymbol{r} \in \mathbb{R}^3\,, \tag{2.2}$$

where the *dielectric permittivity of vacuum* ε_0 defines the proportionality constant required to express the electrostatic field in SI units, $\varepsilon_0 = 8.85418e^{-12}\frac{C}{Vm}$.

Let V be an arbitrary volume in $\mathbb{R}^3$ and ∂V its surface. Further, let F denote an arbitrary, open surface in $\mathbb{R}^3$ and ∂F its boundary as shown in Fig. 2.1. With these notations, the application of Gauss' and Stokes law casts Eqs. (2.1) and (2.2) into the following integral form [61, 76]

$$\forall\, V \subset \mathbb{R}^3,\, \dim V = 3: \qquad \int\limits_{\partial V} d\boldsymbol{f} \cdot \boldsymbol{E} = \frac{1}{\varepsilon_0} \int\limits_{V} \rho(\boldsymbol{r}) d\boldsymbol{r} \tag{2.3}$$

$$\forall\, F \subset \mathbb{R}^3,\, \dim F = 2: \qquad \int\limits_{\partial F} d\boldsymbol{l} \cdot \boldsymbol{E} = 0\,. \tag{2.4}$$

Eq. (2.3) is the so called physical Gauss law. It tells us that every charge distribution ρ in space creates an electrostatic field $\boldsymbol{E}$. This means that charges are the *sources* of the electrostatic field or - in other words - they react to a given electrostatic field. The physical meaning of Eq. (2.2) can be seen in integral form, Eq. (2.4): if we add up on a closed path $\partial F \subset \mathbb{R}^3$ the electric field component parallel to this path, the result is zero. This means that the electrostatic field is free of vortices everywhere.

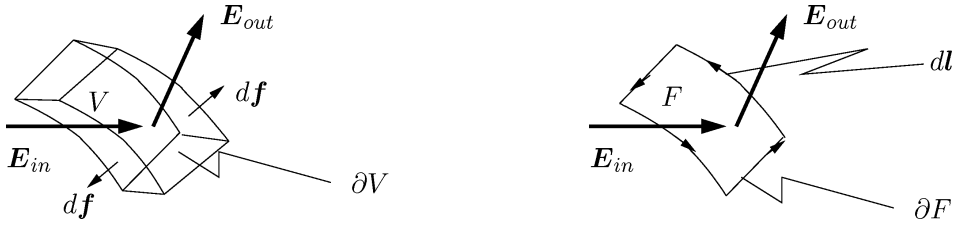


Fig. 2.1: The integral representation of the Maxwell equations.

A vector field fulfilling Eq. (2.2) can be expressed by the gradient of a scalar function, a so called *potential* [29, 61]. Thus, we can introduce the *electrostatic potential* ϕ of the electrostatic field $\boldsymbol{E}$:

$$\boldsymbol{E}(\boldsymbol{r}) = -\boldsymbol{\nabla}\phi(\boldsymbol{r}), \quad \boldsymbol{r} \in \mathbb{R}^3 \tag{2.5}$$

The electrostatic potential ϕ is uniquely defined up to a gauge constant, which does not bear any physical meaning. The constant is commonly chosen in a way that

$$\lim_{\boldsymbol{r}\to\infty} \phi(\boldsymbol{r}) = 0\,.$$

Furthermore, we want to find *physical* solutions $\boldsymbol{E}$ of Eq. (2.1), i.e., solutions for which the field energy W_{field} of the system is finite. Mathematically, this is the case when all the fields appearing in the electrostatic theory (such as the electrostatic field $\boldsymbol{E}$ as well as the dielectric field $\boldsymbol{D}$, which we will define in Section 2.2) are square integrable, i.e., elements of $L^2(\mathbb{R}^3)$. Loosely speaking, this guarantees that the fields in the electrostatic theory vanish "fast enough" for $\boldsymbol{r} \in \mathbb{R}^3$ with $|\boldsymbol{r}| \to \infty$:

$$\lim_{|\boldsymbol{r}|\to\infty} \boldsymbol{E} = \boldsymbol{0}$$

In the following, we call this condition the *radiation condition*.

Inserting Eq. (2.5) into Eq. (2.2) and solving the differential equation

$$\triangle\phi(\boldsymbol{r}) = -\frac{1}{\varepsilon_0}\rho(\boldsymbol{r}), \quad \boldsymbol{r} \in \mathbb{R}^3 \tag{2.6}$$

is then the basic problem of electrostatics [61, 97]. Eq. (2.6) is called *Poisson equation* and the differential operator therein, $\triangle := (\partial_{xx}+\partial_{yy}+\partial_{zz})$ with $\partial_{ii} := \partial_i^2$ for $i \in \{x, y, z\}$, is known as the *Laplace operator*. The potential ansatz reduces the system of first order differential equations (2.1-2) with the vectorial unknown $\boldsymbol{E}$, to a second order differential equation for the unknown, scalar field ϕ. It is often simpler to calculate the potential first and derive $\boldsymbol{E}$ from Eq. (2.5) than to solve for $\boldsymbol{E}$ directly. Thus, we focus on the potential formulation in the following.

The validity of Eqs. (2.1) and (2.2) enforces the electrostatic field $\boldsymbol{E}$ to fulfill continuity conditions at every point $\boldsymbol{r}$ in space. These continuity conditions are now deduced: in order to analyze the behavior of the electrostatic field $\boldsymbol{E}$ on an arbitrary surface $\Gamma \subset \mathbb{R}^3$, we consider the integral forms of the Maxwell equations, Eqs. (2.3) and (2.4).

$$\forall\, V \subset \mathbb{R}^3,\, \dim V = 3: \qquad \int_{\partial V} d\boldsymbol{f} \cdot \boldsymbol{E} = \frac{1}{\varepsilon_0} \int_V \rho(\boldsymbol{r}) d\boldsymbol{r}$$

$$\forall\, F \subset \mathbb{R}^3,\, \dim F = 2: \qquad \int_{\partial F} d\boldsymbol{l} \cdot \boldsymbol{E} = 0\,.$$

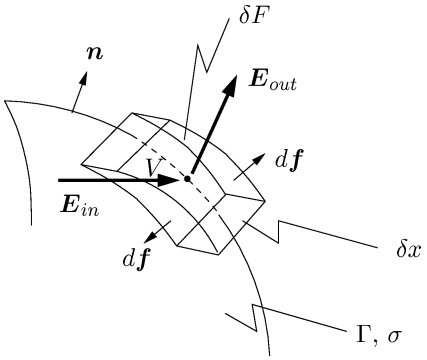


Fig. 2.2: Normal transmission conditions of the electrostatic field (in vacuum).

First, we focus on an infinitesimal small volume V, which comprises a part of Γ as shown in Fig. 2.2. To find the behavior of the electrostatic field $\boldsymbol{E}$ on Γ we perform the limiting process ($\delta x \to 0$) in Eq. (2.3), in a way that δV contracts to the small, black point drawn in Fig. 2.2: the *left hand side* yields

$$\lim_{\delta x \to 0} \int\limits_{\partial V} \delta \boldsymbol{f} \cdot \boldsymbol{E} = \boldsymbol{n} \cdot (\boldsymbol{E}_{out} - \boldsymbol{E}_{in})\, \delta F\,, \tag{2.7}$$

where $\boldsymbol{n}\,\delta F = \delta \boldsymbol{f}$. In Eq. (2.7) we used the fact that δF is infinitesimal small and this allows to assume that the electrostatic fields $\boldsymbol{E}_{in}$ and $\boldsymbol{E}_{out}$ are constant. On the *right hand side* we find a non-vanishing result for the limiting process only if the surface Γ has a singular surface charge σ, i.e., if the charge is distributed solely on Γ:

$$\lim_{\delta x \to 0} \int\limits_{V} d\boldsymbol{r} \frac{\rho(\boldsymbol{r})}{\varepsilon_0} = \frac{\sigma}{\varepsilon_0} \delta F\,. \tag{2.8}$$

Combining Eqs. (2.7) and (2.8), the limiting process results in

$$\boldsymbol{n} \cdot (\boldsymbol{E}_{out} - \boldsymbol{E}_{in}) = \frac{\sigma}{\varepsilon_0}\,. \tag{2.9}$$

This means that the normal component of the electrostatic field, $(\boldsymbol{n} \cdot \boldsymbol{E})$, is discontinuous on a surface that carries a charge density σ. For $\sigma = 0$, the normal component is continuous.

The behavior of the tangential components of $\boldsymbol{E}$ becomes clear by considering Eq. (2.4) in the limiting process ($\delta x \to 0$) and infinitesimal small $\delta l = |\delta \boldsymbol{l}| \neq 0$ (see Fig. 2.3 for notations)

$$\begin{aligned} 0 &= \lim_{\delta x \to 0} \int\limits_{\partial F} \delta \boldsymbol{l} \cdot \boldsymbol{E} \\ \Leftrightarrow \quad 0 &= \delta \boldsymbol{l} \cdot (\boldsymbol{E}_{out} - \boldsymbol{E}_{in}) = \underbrace{\delta l\, (\boldsymbol{t} \times \boldsymbol{n})}_{\delta \boldsymbol{l}} \cdot (\boldsymbol{E}_{out} - \boldsymbol{E}_{in}) \\ \Leftrightarrow \quad 0 &= (\boldsymbol{t} \times \boldsymbol{n}) \cdot (\boldsymbol{E}_{out} - \boldsymbol{E}_{in})\,. \end{aligned}$$

We introduced $\boldsymbol{n}$ and $\boldsymbol{t}$, which, together with $\boldsymbol{l}$, complete the orthogonal trihedral on Γ. The vector $\boldsymbol{t}$ points out of the drawing plane and is tangential to the surface Γ.

Eq. (2.1) is valid for an arbitrary setup enclosing Γ, i.e., for every vector $\boldsymbol{t}$ tangential to the surface. Thus, we can generalize

$$\boldsymbol{n} \times (\boldsymbol{E}_{out} - \boldsymbol{E}_{in}) = \boldsymbol{0}\,. \tag{2.10}$$

The continuity of the tangential component of the electrostatic field, Eq. (2.10), transfers to the

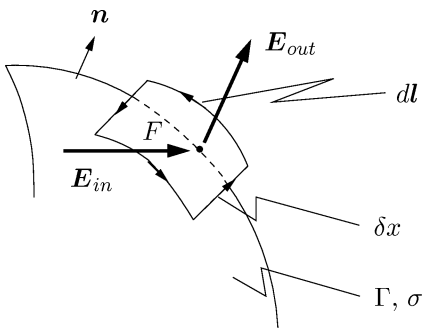


Fig. 2.3: Tangential transmission conditions of the electrostatic field.

continuity of the electrostatic field ϕ at every arbitrary point $\boldsymbol{r} \in \mathbb{R}^3$ [29], see Appendix 10.2 where we recall the proof given in [52]:

$$\phi_{out} - \phi_{in} = 0$$

Our findings on the transmission conditions of the electrostatic field $\boldsymbol{E}$ are subsumed in Theorem 2.1.1.

Theorem 2.1.1 *Maxwell equations in vacuum and transmission conditions: Let us assume a static, and therefore fixed charge density, ρ in $\mathbb{R}^3$. Additionally, let us assume a static surface charge distribution σ spread over an arbitrary surface $\partial V \subset \mathbb{R}^3$.*
The Maxwell equations in differential form read

$$\begin{aligned} \boldsymbol{\nabla} \cdot \boldsymbol{E} &= \frac{1}{\varepsilon_0}\rho, && \text{in } \mathbb{R}^3 \\ \boldsymbol{\nabla} \times \boldsymbol{E} &= \boldsymbol{0}, && \text{in } \mathbb{R}^3\,. \end{aligned}$$

Due to the surface charge density σ, the electrostatic field $\boldsymbol{E}$ fulfills the following transmission conditions on ∂V

$$\begin{aligned} \boldsymbol{n} \times (\boldsymbol{E}_{out} - \boldsymbol{E}_{in}) &= \boldsymbol{0}, && \text{on } \partial V \\ \boldsymbol{n} \cdot (\boldsymbol{E}_{out} - \boldsymbol{E}_{in}) &= \sigma(\boldsymbol{r}), && \text{on } \partial V\,. \end{aligned}$$

We close this paragraph with an important remark on the Maxwell equations, which is used several times in this work.:

Important remark 2.1 The Maxwell equations are *linear* equations. Therefore, the principle of superposition is valid [61]. This means that the net electrostatic field at a given place caused by two or more charges is the sum of the electrostatic fields which would have been caused by each charge individually.

2.1.1 The Laplace operator

In the above discussion, we assumed that Eq. (2.6) is valid in $\mathbb{R}^3$ and further that the radiation condition has to be fulfilled by any physical solution. In this case, the solution of Eq. (2.6) can be analytically written in the form of an integral. In order to show this, we introduce the so called *fundamental solution* of the Laplace operator.

Theorem 2.1.2 *The fundamental solution of the Laplace operator: A solution to the following differential equation*[1] *defined in* $\mathbb{R}^3$

$$\triangle G_{\triangle}(\boldsymbol{r}) = -\delta(\boldsymbol{r}), \quad \boldsymbol{r} \in \mathbb{R}^3 \tag{2.11}$$

is given by

$$\begin{aligned} G_{\triangle}: \quad \mathbb{R}^3\backslash\{0\} &\mapsto \mathbb{R} \\ \boldsymbol{r} &\mapsto \frac{1}{4\pi\,|\boldsymbol{r}|} \end{aligned} \tag{2.12}$$

$G_{\triangle}$ is called the fundamental solution of the Laplace operator $\triangle$. $G_{\triangle}$ is the unique, physical solution to Eq. (2.11)*, i.e., when the radiation condition has to be fulfilled.*
Proof: see [97, 121].

With the knowledge of the fundamental solution of the Laplace operator, it is easy to find the solution to problem (2.6). The electrostatic potential ϕ is proportional to the following convolution:

$$f(\boldsymbol{r}) := (G_{\triangle} * \rho)(\boldsymbol{r}) := \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, G_{\triangle}(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r}'), \quad \boldsymbol{r} \in \mathbb{R}^3\,, \tag{2.13}$$

as

$$\triangle f(\boldsymbol{r}) = \triangle[(G_{\triangle} * \rho)(\boldsymbol{r})] = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, \underbrace{\triangle G_{\triangle}(\boldsymbol{r}-\boldsymbol{r}')}_{-\delta(\boldsymbol{r}-\boldsymbol{r}')}\rho(\boldsymbol{r}') = -\rho(\boldsymbol{r}), \quad \boldsymbol{r} \in \mathbb{R}^3\,. \tag{2.14}$$

The function f defined in Eq. (2.13) is called the *Newton potential* of the Laplace operator with source term ρ [27, 52, 121].

Often, the validity of Eq.(2.6) is *not* given in the whole of $\mathbb{R}^3$, but is confined to a *finite* volume $V \subset \mathbb{R}^3$, where we are given fixed values of the potential ϕ (Dirichlet data) or its normal derivative at the volume's boundary ∂V (Neumann data).

For these so called *boundary value* or transmission problems, Eq. (2.13) is obviously no longer valid. It *has* to be modified in order to fulfill the boundary values. The idea behind this is to add a correction term to the Newton potential, which does not impact the differential equation. Mathematically spoken, the correction term lies in the kernel of the respective operator and it modifies the Newton potential so that the boundary values are correct [2].

For instance, assume to have the following Dirichlet boundary value problem

$$\left.\begin{aligned} \triangle\phi(\boldsymbol{r}) &= \delta(\boldsymbol{r}), \quad \boldsymbol{r} \in V \\ \phi(\boldsymbol{r}) &= b(\boldsymbol{r}), \quad \boldsymbol{r} \in \partial V\,. \end{aligned}\right\} \tag{2.15}$$

Let $f_0 \in \mathrm{Kern}(\triangle)$, i.e. $\triangle f_0 = 0$. The solution of system (2.15) is then given by

$$\phi(\boldsymbol{r}) = (G_{\triangle} * \rho)(\boldsymbol{r}) + f_0(\boldsymbol{r}), \quad \boldsymbol{r} \in V,$$

where f_0 is chosen in a way that the boundary condition is fulfilled.

[1] Eq. (2.11) has to be understood in the distributional sense, because the *right hand side* is a δ-distribution.
[2] Compare (a) low-level mathematics where we search for general homogeneous and special solutions to solve differential equations (b) the "weak" formulation for the inhomogeneous solution of the Yukawa operator in Section 4.3 (Single and Double layer potentials) and (c) the Helmholtz decomposition derived in [29].

It is, for instance, easy to show that the kernel of the Laplace operator for spherically symmetric problems, i.e., the solution of the differential equation

$$\triangle f_0(r) = 0, \qquad r = |\boldsymbol{r}|, \boldsymbol{r} \in \mathbb{R}^3 \,,$$

is given by

$$f_0(r) = \begin{cases} A + \frac{B}{r}, & r \neq 0 \\ A, & r = 0 \end{cases}$$

where $A, B \in \mathbb{R}$ are constants.

2.2 The Maxwell equations in dense media

Dense media consist of so many atoms that an explicit treatment in terms of their electrostatics is difficult. The macroscopic material equations offer a possibility to account for their reaction, as they try to capture the electrostatic effect of all bound charges of the medium's molecules within a *mean-field* approach. In order to motivate this and to set up the basis for the macroscopic theory, we discuss the multipole series of an arbitrary charge distribution in Section 2.2.1. Then, in Section 2.2.2 we introduce the material equations which are derived from the Maxwell equations in vacuum. Section 2.2.3 finally deals with the linear response theory which is required to specify the medium's reaction to an external electrostatic field.

2.2.1 Monopole and dipole moment of a charge distribution

In Section 2.1.1 we learned that the solution to Eq. (2.6) can be written as

$$\phi(\boldsymbol{r}) = \frac{1}{\varepsilon_0}(G_\triangle * \rho)(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0} \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' \frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}, \quad \boldsymbol{r} \in \mathbb{R}^3 \,. \tag{2.16}$$

The calculation of the volume integral is often complicated. However, if one is interested in the potential $\phi(\boldsymbol{r})$ "far" from the charge distribution as illustrated in Fig. 2.4, it often suffices to develop $\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}$ into a Taylor series around the origin, i.e., $\boldsymbol{r}' = \boldsymbol{0}$ [61]. The Taylor series of Eq. (2.16) is given by

$$4\pi\varepsilon_0\phi(\boldsymbol{r}) = \frac{Q}{r} - (\boldsymbol{\nabla}\frac{1}{r}) \cdot \boldsymbol{p} + \cdots = \frac{Q}{r} + \frac{\boldsymbol{r} \cdot \boldsymbol{p}}{r^3} + \mathcal{O}(\frac{1}{r^5}) \,, \quad r = |\boldsymbol{r}|, \boldsymbol{r} \in \mathbb{R}^3 \,, \tag{2.17}$$

where we used the following definitions:

Definition 2.1 Monopole and dipole:

$$\begin{aligned} Q &:= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' \rho(\boldsymbol{r}') \\ \boldsymbol{p} &:= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' \boldsymbol{r}' \rho(\boldsymbol{r}') \end{aligned}$$

Q denotes the monopole or total charge and $\boldsymbol{p}$ is known as dipole moment.

We truncate the Taylor series after the dipole moment, as the higher moments decrease on the order $\geq \mathcal{O}(\frac{1}{r^5})$ and therefore, they are often negligible compared to the monopole and dipole moment. The approximation then reads

$$\phi(\boldsymbol{r}) = \frac{Q}{4\pi\varepsilon_0\, r} + \frac{\boldsymbol{r} \cdot \boldsymbol{p}}{4\pi\varepsilon_0\, r^3} \,, \quad r = |\boldsymbol{r}|, \boldsymbol{r} \in \mathbb{R}^3 \,. \tag{2.18}$$

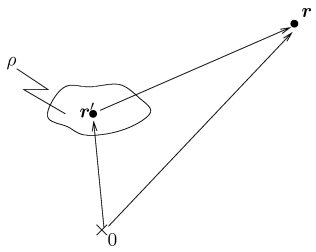


Fig. 2.4: A Taylor series around $\boldsymbol{r}' = 0$ is used to approximate the potential at position $\boldsymbol{r}$ far from the charge distribution ρ.

When the total charge is zero, i.e., $Q = 0$, the dipole moment dominates. It is a measure for the symmetry of the charge distribution. From this point of view the multipole series does not only describe the asymptotic behavior of the electrostatic potential, but gives an idea of the spatial distribution and the complexity of ρ. Generally, the more moments are taken into account the better the "resolution" of the charge distribution.

Since the monopole moment does not provide any hint on the spatial distribution but only on the overall magnitude of ρ, we will have a closer look at the dipole moment. The arrangement of two opposite point charges with magnitude q and a distance vector $\boldsymbol{a}$ from the negative to the positive charge is called a *dipole*. The Taylor series of this charge distribution yields an expression for the dipole moment, namely

$$\boldsymbol{p} = q\boldsymbol{a}\,.$$

Fig. 2.5 illustrates a dipole within an external electrostatic field $\boldsymbol{E}$. Due to the orientation of the dipole, a torque $\boldsymbol{N}$ acts on the dipole

$$\boldsymbol{N} = \boldsymbol{p} \times \boldsymbol{E}\,.$$

Because of this torque, the dipole tries to align parallel to the electrostatic field $\boldsymbol{E}$ to minimize its potential energy.

If we assume a high number of dipoles between the two plates of the capacitor in Fig. 2.5, all dipoles try to align parallel to the external electric field and the total electrostatic field decreases in its magnitude due to the sum of opposing dipole fields [61]. Such a weakening of the electrostatic field is called screening effect in the following.

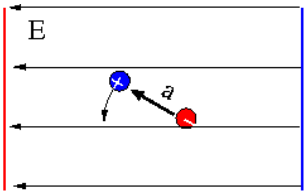


Fig. 2.5: An electric dipole orients parallel to the external field.

In the following section we introduce the polarization field $\boldsymbol{P}$ as a measure of the effective dipole moment density of a dense medium. As the occurrence of a dipole moment $\boldsymbol{p}$ in a dense medium can have very different sources, we distinguish between different kinds of polarization [61]: the deformations of the protons and electrons in the atoms create a weak *electronic* or *deformation polarization*. A permanent dipole moment of the medium's molecules can cause a huge polarization in presence of an external electrostatic field because of their orientation, see the discussion above and Section 3.2.1. A polarization due to the orientation of permanent dipoles is called *Debye* or

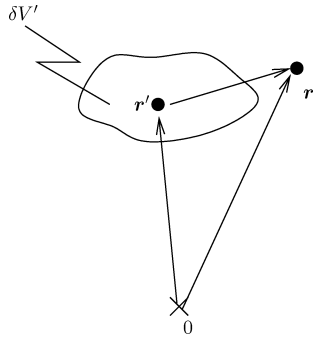


Fig. 2.6: A Taylor series around $\boldsymbol{r}' \in \delta V'$ is used to describe the influence of the medium in $\delta V'$ at position $\boldsymbol{r}$.

orientation polarization [19]. Polarization effects which occur in a dielectric medium are often subsumed under the term *dielectric response* of the medium.

2.2.2 The macroscopic material equations

In principle, the Maxwell equations given in Section 2.1,

$$\begin{aligned} \boldsymbol{\nabla} \cdot \boldsymbol{E}(\boldsymbol{r}) &= \frac{1}{\varepsilon_0}\rho(\boldsymbol{r}), & \boldsymbol{r} &\in \mathbb{R}^3 \\ \boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}) &= \boldsymbol{0}, & \boldsymbol{r} &\in \mathbb{R}^3\,, \end{aligned}$$

suffice to describe the electrostatic effect of an arbitrary static charge distribution. It requires to define *every* charge in the system as a fixed charge: not only the excess partial charges of the biomolecule, but also all bound charges in every atom of the system.

However, there are two crucial problems, which cannot be directly solved: the huge number of solvent molecules ($\sim 10^{23}$) makes an explicit modeling impossible – at least in the framework of the Maxwell equations. The other problem is that neither the solvent molecules nor their bound charges are fixed because of the thermal energy of the molecules and the quantum mechanical nature of their electrons.

Even though this seems a problem we cannot cope with, it turns out that many experiments measure macroscopic, thermodynamic quantities of the medium and thus, it is often sufficient to search for an appropriate macroscopic theory, within which the microscopic behavior is locally averaged in space and time. In order to deduce a macroscopic theory of electrostatics for a dense medium, let us assume that the sum over all bound charges of an atom is zero, otherwise they would have been assigned to the fixed charge distribution ρ. Then, the dominant contribution of the bound charge distribution is the dipole moment. We define the macroscopic electrostatic polarization $\boldsymbol{P}$ at $\boldsymbol{r}' \in \mathbb{R}^3$ by the average over all dipoles in a small volume $\delta V'$ around $\boldsymbol{r}'$

$$\boldsymbol{P}(\boldsymbol{r}') = \sum_{i \in \delta V'} N_i \langle \boldsymbol{p}_i \rangle\,,$$

where the average dipole $\langle \boldsymbol{p}_i \rangle$ of atom type i is weighted by the number N_i of atoms of type i found in $\delta V'$.

A Taylor series around $\boldsymbol{r}' \in \delta V'$ yields the potential at position $\boldsymbol{r}$ that is induced by the fixed excess charges and the microscopic dipoles in this volume. Due to Eq. (2.18) on p. 14 we can write

$$\delta\phi(\boldsymbol{r}, \boldsymbol{r}') \;=\; \frac{1}{4\pi\varepsilon_0}\left[\frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} + \boldsymbol{P}(\boldsymbol{r}') \cdot \left(\boldsymbol{\nabla}_{r'} \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right)\right] \delta V'\,.$$

The volume $\delta V'$ is macroscopically so small that we replace it by $d\boldsymbol{r}'$ as well as $\delta\phi(\boldsymbol{r},\boldsymbol{r}')$ by $d\phi(\boldsymbol{r},\boldsymbol{r}')$. An integration by parts yields

$$\begin{aligned}\phi(\boldsymbol{r}) &= \frac{1}{4\pi\varepsilon_0}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\left[\frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}+\boldsymbol{P}(\boldsymbol{r}')\cdot\left(\boldsymbol{\nabla}_{r'}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\right)\right]\\ &= \frac{1}{4\pi\varepsilon_0}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\left[\frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}-\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\left(\boldsymbol{\nabla}_{r'}\cdot\boldsymbol{P}(\boldsymbol{r}')\right)\right]\\ &= \frac{1}{\varepsilon_0}[G_\triangle * (\rho-\boldsymbol{\nabla}\cdot\boldsymbol{P})(\boldsymbol{r})]\,.\end{aligned}$$

The application of the Laplace operator and further using Eq. (2.14)[3] finally give the important result

$$\boldsymbol{\nabla}\cdot(\varepsilon_0\boldsymbol{E}(\boldsymbol{r})+\boldsymbol{P}(\boldsymbol{r}))=\rho(\boldsymbol{r}),\quad \boldsymbol{r}\in\mathbb{R}^3\,. \tag{2.19}$$

In Eq. (2.19) the expression $(-\boldsymbol{\nabla}\cdot\boldsymbol{P})$ can be interpreted as an effective charge distribution ρ_{ind}. The additional source ρ_{ind} vanishes, if the divergence of the polarization is zero. This is reasonable, because, when representing the medium in a small volume $\delta V'$ as a superposition of the dipoles contained in $\delta V'$, we have a net source only if the dipoles sum up to a local excess charge[4].

Further, if we interpret the reaction of the medium as a process to minimize the potential energy of the system, it is reasonable to assume that the polarization weakens the resulting electrostatic field $\boldsymbol{E}$.

Eq. (2.19) motivates the definition of the so called *dielectric (displacement) field* or *dielectric permittivity*

$$\boldsymbol{D} := (\varepsilon_0\boldsymbol{E}+\boldsymbol{P})\,.$$

The dielectric permittivity $\boldsymbol{D}$ is only affected by the fixed charges,

$$\boldsymbol{\nabla}\cdot\boldsymbol{D}=\rho\,, \tag{2.20}$$

which means that in dielectric media the boundary condition for the normal component of the electrostatic field, $\boldsymbol{n}\cdot\boldsymbol{E}$, in vacuum is valid for $\boldsymbol{D}$ in dense media, as Eq. (2.20) formally corresponds to Eq. (2.1). This implies that for vanishing surface charge distribution, $\sigma=0$, the normal component of the dielectric permittivity, $\boldsymbol{n}\cdot\boldsymbol{D}$ is continuous. However, the normal component of the electric field, $\boldsymbol{n}\cdot\boldsymbol{E}$ is not:

$$\begin{aligned} & \sigma(\boldsymbol{r}) = 0, && \boldsymbol{r}\in\Gamma\\ \Leftrightarrow\quad & \boldsymbol{n}\cdot(\boldsymbol{D}_{out}-\boldsymbol{D}_{in}) = 0, && \boldsymbol{r}\in\Gamma\\ \Leftrightarrow\quad & \boldsymbol{n}\cdot(\boldsymbol{E}_{out}-\boldsymbol{E}_{in}) = -\boldsymbol{n}\cdot(\boldsymbol{P}_{out}-\boldsymbol{P}_{in}), && \boldsymbol{r}\in\Gamma\end{aligned}$$

This means that a change of the polarization $\boldsymbol{P}$ on a common boundary $\Gamma\subset\mathbb{R}^3$ causes a jump in the normal component and can therefore be interpreted as an induced surface charge σ_{ind}[5].

In contrast to Eq. (2.1), which has to be corrected by the polarization field $\boldsymbol{P}$ in order to account for a dielectric medium, Eq. (2.2),

$$\boldsymbol{\nabla}\times\boldsymbol{E}=\boldsymbol{0}\,,$$

[3] Using Eq. (2.14) implies that the electrostatic field fulfills the radiation condition.

[4] This is, for instance, the case on boundaries where the composition of the material is changed and motivates σ_{ind} to denote *induced surface charge*.

[5] Compare footnote 4.

does not change for dense media. This implies that the tangential component of the electrostatic field is continuous. We state the electrostatic material equations.

Theorem 2.2.1 *The Maxwell equations in dielectric media: We consider the dielectric medium to be spread out in $\mathbb{R}^3$. We define its polarization field $\boldsymbol{P}$. A charge density ρ is embedded in the medium.*

The material equations determining the total electrostatic field $\boldsymbol{E}$ in $\mathbb{R}^3$ read

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{r}) &= \varepsilon_0 \boldsymbol{E}(\boldsymbol{r}) + \boldsymbol{P}(\boldsymbol{r}) && (2.21)\\ \boldsymbol{\nabla} \cdot \boldsymbol{D}(\boldsymbol{r}) &= \rho(\boldsymbol{r}) && (2.22)\\ \boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r}) &= \boldsymbol{0}\,, && (2.23) \end{aligned}$$

for $\boldsymbol{r} \in \mathbb{R}^3$.

Consider an arbitrary surface $\partial V \subset \mathbb{R}^3$ as it is shown in Figs. 2.2 and 2.3 for the vacuum case. ∂V has a surface charge density σ. Then, the dielectric and the electric field fulfill the following transmission conditions on ∂V

$$\begin{aligned} \boldsymbol{n} \times (\boldsymbol{E}_{out}(\boldsymbol{r}) - \boldsymbol{E}_{in}(\boldsymbol{r})) &= 0\\ \boldsymbol{n} \cdot (\boldsymbol{D}_{out}(\boldsymbol{r}) - \boldsymbol{D}_{in}(\boldsymbol{r})) &= \sigma(\boldsymbol{r})\,, \end{aligned}$$

with $\boldsymbol{r} \in \partial V$.

2.2.3 The dielectric operator

In the preceding section, we discussed that the medium is an additional charge carrier, which reacts to the presence of fixed excess charges ρ. We speak of a *dielectric response* of the medium. The response has been taken to originate from the dipole moment of the bound charges of the medium and is measured by the macroscopic polarization field $\boldsymbol{P}$.

This section concerns the question *how* $\boldsymbol{E}$, $\boldsymbol{P}$, and $\boldsymbol{D}$ depend on each other. We introduce a general concept for describing the dielectric response within a *linear* response theory [74].

Having in mind the picture of small dipoles rotating in the external field $\boldsymbol{D}$ to minimize their potential energy, it is reasonable to assume that the magnitude of the polarization $\boldsymbol{P}$ *depends* on the field that is caused by the fixed charges

$$\boldsymbol{P} = \boldsymbol{P}(\boldsymbol{D})\,.$$

Using Eq. (2.21) in turn gives rise to the following functional dependencies

$$\boldsymbol{D} = \boldsymbol{D}(\boldsymbol{E}) \quad \text{and} \quad \boldsymbol{P} = \boldsymbol{P}(\boldsymbol{E})\,.$$

A linear response theory predicts an increase in the response (in our case the polarization field $\boldsymbol{P}$), when the external driving field (in our case the dielectric field $\boldsymbol{D}$) gets stronger.

However, thinking of the dielectric response, which is caused by the dipole reorientation, there is a point where a further increase in magnitude of the dielectric field will not change the polarization anymore, since all dipoles are already aligned. Such a *saturation* can only be captured by a nonlinear theory. Since the saturation of dipole alignment does not occur until there are high electrostatic fields, they can be taken as an effect of higher order [19, 124]. This justifies a *linear* response - at least in order to study the overall influence of the polarization effects we are primarily interested in.

The simplest model of a linear dielectric response is given by

$$\boldsymbol{P}(\boldsymbol{r}) = \chi(\boldsymbol{r})\,\boldsymbol{D}(\boldsymbol{r}), \qquad \boldsymbol{r} \in \mathbb{R}^3 \tag{2.24a}$$

$$\boldsymbol{P}(\boldsymbol{r}) = (\varepsilon(\boldsymbol{r}) - \varepsilon_0)\,\boldsymbol{E}(\boldsymbol{r}), \qquad \boldsymbol{r} \in \mathbb{R}^3 \tag{2.24b}$$

$$\boldsymbol{D}(\boldsymbol{r}) = \varepsilon(\boldsymbol{r})\,\boldsymbol{E}(\boldsymbol{r}), \qquad \boldsymbol{r} \in \mathbb{R}^3\,. \tag{2.24c}$$

Here, the macroscopic state of the system is described by the so called *dielectric function* $\varepsilon(\boldsymbol{r})$ measured in units of ε_0 and the *dielectric response function* (or dielectric susceptibility) $\chi(\boldsymbol{r}) = (1-\varepsilon^{-1})(\boldsymbol{r})$ [6].

Whether such an approximation is justified or not depends on the medium under consideration. In Eqs. (2.24)(a-c) the reaction is assumed to have neither spatial nor temporal dependencies: it is a *local* reaction. This means that at an arbitrary position $\boldsymbol{r} \in \mathbb{R}^3$ where the fixed charges create the field $\boldsymbol{D}(\boldsymbol{r})$, the bound charges of the molecules in a small volume around $\boldsymbol{r}$ align with this field without any influence of the neighboring molecules.

The alignment leads to a further contribution to the electrostatic field $\boldsymbol{E}(\boldsymbol{r})$ in a way that it locally weakens the electrostatic field. Assuming Eqs. (2.24)(a-c) to be valid, this macroscopic screening is given by the factor ε. Since the dielectric function ε ranges between $1\varepsilon_0$ and $100\varepsilon_0$ for commonly used solvents, the effect of the medium on the electrostatic field is very important [93, 133]. Based on models for the polarization (Langevin dipole, Onsager and Debye model) one can calculate the local dielectric constant for polar and non-polar continua [19].

The mean field approach with constants, ε and χ, lacks in modeling the correlations or interactions of the medium's molecules. Such interactions, however, can lead to a measurable change of the dielectric response. To account for spatial deviations from the bulk dielectric response and in order to account for the dependency on all previous time steps, $t' < t$, Eqs. (2.24)(a-c) have to be extended to a more general form, see [33, 74, 124]:

$$P_i(t,\boldsymbol{r}) = \int_{-\infty}^{t}\int_{\mathbb{R}^3} \left(\varepsilon_{ij}(t-\tau;\boldsymbol{r},\boldsymbol{r}') - \varepsilon_0\delta_{ij}\right) E_j(\tau,\boldsymbol{r}')d\boldsymbol{r}'d\tau, \qquad \text{in } \mathbb{R}^3 \tag{2.25a}$$

$$P_i(t,\boldsymbol{r}) = \int_{-\infty}^{t}\int_{\mathbb{R}^3} \chi_{ij}(t-\tau;\boldsymbol{r},\boldsymbol{r}')D_j(\tau,\boldsymbol{r}')d\boldsymbol{r}'d\tau, \qquad \text{in } \mathbb{R}^3 \tag{2.25b}$$

$$D_i(t,\boldsymbol{r}) = \int_{-\infty}^{t}\int_{\mathbb{R}^3} \varepsilon_{ij}(t-\tau;\boldsymbol{r},\boldsymbol{r}')E_j(\tau,\boldsymbol{r}')d\boldsymbol{r}'d\tau, \qquad \text{in } \mathbb{R}^3\,, \tag{2.25c}$$

where we used the Einstein notation [7].

The constants, χ and ε, in Eqs. (2.24)(a-c) are replaced by tensors of second order. Eqs. (2.25)(a-c) account for spatial dispersion meaning that the response at position $\boldsymbol{r}$ depends on every surrounding position $\boldsymbol{r}'$. In contrast to the local response captured by the Eqs. (2.24)(a-c), this behavior is of a *nonlocal* nature. Such a dependency originates from correlations of nearby molecules and the local variations of the electrostatic field $\boldsymbol{E}$.

In a common dielectric, the spatial dispersion is not an issue since the integral kernel decays on distances $|\boldsymbol{r}-\boldsymbol{r}'|$ smaller than an atomic dimension [74]. In this case, we can assume that the

[6] For more general ε, ε^{-1} turns out to be the inverse operator; χ should not be confused with the electric susceptibility $\chi_e(\boldsymbol{r}) := \varepsilon(\boldsymbol{r}) - 1$.

[7] When an index variable appears twice in a single term, it implies that we are summing over all of its possible values {1,2,3}.

macroscopic averaged field $\boldsymbol{E}$ has small variations, namely

$$\boldsymbol{E}(\boldsymbol{r}') \sim \boldsymbol{E}(\boldsymbol{r})\,.$$

Then, $\boldsymbol{E}$ can be extracted from the integral and we return to Eqs. (2.24)(a-c). However, when the length scale of the correlations exceeds the atomic scale, the dispersion yields new physical effects which have to be considered. For example, this is the case for conductors and electrolytes [74], but also for the water molecules with their hydrogen bonds and permanent dipoles [33]. Thus, in Section 3.2 we develop a nonlocal, linear dielectric operator for water.

2.3 Coulomb interaction energy and field energy

In the following, we recapitulate the expressions for the electrostatic (Coulomb) interaction energy W_{coul} and the electrostatic field energy W_{field} in a dielectric medium. It is the basis for the considerations in Section 3.1.4, where we describe the change in free energy $\triangle G_{solv}$ when the molecule is transferred from vacuum (or the gas phase) to the solvent. From the experimental point of view, $\triangle G_{solv}$ is directly measurable and therefore serves as a quantity to assess and optimize a theoretical model.

The electrostatic energy W_{coul} of a test charge q at position $\boldsymbol{r}_q$ in an external, electrostatic field $\boldsymbol{E} = -\boldsymbol{\nabla}\phi$, is given by

$$W_{coul} = q\,\phi(\boldsymbol{r}_q)\,. \tag{2.26}$$

It equals the work that has to be done to move the test charge from infinity to its position $\boldsymbol{r}_q$ in the electrostatic field [61]. This energy is called *Coulomb interaction energy*. The Coulomb interaction energy of an arbitrary charge distribution

$$\rho(\boldsymbol{r}) = \sum_i q_i \delta(\boldsymbol{r} - \boldsymbol{r}_i) \tag{2.27}$$

is given by

$$W_{coul} = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\rho(\boldsymbol{r})\phi(\boldsymbol{r}) = \sum_i q_i\,\phi(\boldsymbol{r}_i)\,. \tag{2.28}$$

Please note that we made no further assumption on the origin of the potential ϕ in Eq. (2.28) meaning that the dielectric surrounding and the fixed sources are not specified, but reflected in $\phi(\boldsymbol{r})$. Thus, Eq. (2.28) is valid for an *arbitrary* electrostatic setting.

However, the interaction energy given in Eq. (2.28) does not tell us anything about the energy needed to *create* the field itself. The so called *field energy* W_{field} is defined by the successive build-up of the charge distribution. In particular, in a dielectric medium the polarization $\boldsymbol{P}$ causes an additional energy contribution. To find an expression for the field energy, we start with the infinitesimal work which has to be done to increase the charge distribution from $\rho\,d\boldsymbol{r}$ to $(\rho + \delta\rho)\,d\boldsymbol{r}$ in the potential ϕ created by the already present charge distribution ρ [8]

$$\delta W_{field} = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\,\delta\rho\,\phi\,.$$

[8] Another more general access to the field energy is to express it as a functional of the *independent* fields $\boldsymbol{D}$ and $\boldsymbol{P}$, see [86, 102]. In addition, this approach is easily extensible to account for energy terms resulting from ionic solvents, see [7].

With Eq. (2.22) we substitute $\delta\rho$ by $\delta(\boldsymbol{\nabla}\cdot\boldsymbol{D})$:

$$\begin{aligned}\delta W_{field} &= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\,\delta(\boldsymbol{\nabla}\cdot\boldsymbol{D})\,\phi = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\,\boldsymbol{\nabla}\cdot(\delta\boldsymbol{D})\,\phi \\ &= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}(\delta\boldsymbol{D})\cdot(\boldsymbol{\nabla}\phi) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}(\delta\boldsymbol{D})\cdot\boldsymbol{E}\,.\end{aligned}$$

In this derivation we applied the Gauss law and used the fact that the electrostatic field fulfills the radiation condition. An integration over $\boldsymbol{D}$ finally yields

$$W_{field} = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\left(\int_0^{\boldsymbol{D}}\delta\boldsymbol{D}\cdot\boldsymbol{E}\right)(\boldsymbol{r})\,. \tag{2.29}$$

A further evaluation is only possible if we specify the dielectric reaction of the medium. Assuming, for instance, a constant macroscopic dependence of the dielectric permittivity $\boldsymbol{D}$ and the electrostatic field $\boldsymbol{E}$,

$$\boldsymbol{D} = \varepsilon\boldsymbol{E}\,,$$

as introduced in Section 2.2.2, we can integrate Eq. (2.29) by parts. This yields the following field energy:

$$\begin{aligned}W_{field} &= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\left(\int_0^{\boldsymbol{D}}\delta\boldsymbol{D}\cdot\boldsymbol{E}\right)(\boldsymbol{r}) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\varepsilon\left(\int_0^{\boldsymbol{E}}\delta\boldsymbol{E}\cdot\boldsymbol{E}\right)(\boldsymbol{r}) \\ &= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\varepsilon\left(\int_0^{\boldsymbol{E}}(E_1dE_1+E_2dE_2+E_3dE_3)\right)(\boldsymbol{r}) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\varepsilon\left(\frac{1}{2}(E_1^2+E_2^2+E_3^2)\right)(\boldsymbol{r}) \\ &= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\left(\frac{1}{2}(\varepsilon\boldsymbol{E}\cdot\boldsymbol{E})\right)(\boldsymbol{r}) = \frac{1}{2}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}\boldsymbol{D}(\boldsymbol{r})\cdot\boldsymbol{E}(\boldsymbol{r})\,.\end{aligned} \tag{2.30}$$

The term $\left(\frac{1}{2}\,\boldsymbol{D}\cdot\boldsymbol{E}\right)$ in Eq. (2.30) is usually interpreted as the *field energy density* of the system. Another interpretation of the field energy is revealed when $\boldsymbol{D}$ is once again substituted by the corresponding Maxwell equation:

$$\begin{aligned}W_{field} &= \frac{1}{2}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}\boldsymbol{D}(\boldsymbol{r})\cdot\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{2}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}(\boldsymbol{\nabla}\cdot\boldsymbol{D}(\boldsymbol{r}))\,\phi(\boldsymbol{r}) \\ &= \frac{1}{2}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}\rho(\boldsymbol{r})\,\phi(\boldsymbol{r})\,.\end{aligned} \tag{2.31}$$

In this formulation, it is emphasized that the field energy W_{field} is caused by a particular charge distribution ρ.

Comparing Eq. (2.26) and Eq. (2.31), we note that the field energy is half the interaction energy. The difference originates from the fact that the field energy accounts for the "hypothetical charging process", where the electrostatic field $\boldsymbol{E}$ and therefore the potential ϕ is built *by* the added charges. In contrast, the interaction energy of a charge q with an external electrostatic field $\boldsymbol{E}$ assumes the field to be already built. The charge q does *not* contribute to this external field $\boldsymbol{E}$.

A problem, which we have not commented on yet, is the so called *Coulomb self energy* term

that appears in Eq. (2.31), when assuming the charge distribution to be a set of point charges as in Eq. (2.27). With Theorem 2.1.2, the integration turns into a double summation, which yields contributions of the form

$$\sim \frac{q_i\, q_i}{|\boldsymbol{r}_i - \boldsymbol{r}_i|}\,.$$

These terms take into account the (infinite) energy needed to assemble the point charge q_i itself. In physical applications, however, this infinite contribution cancels out as we will see in Section 3.1.4.

Chapter 3

The biomolecular system

Modeling biomolecules in their natural surrounding, namely in living organisms, is of great importance for a detailed understanding of the diversity of biochemical processes exhibited by the biomolecules. In particular, the electrostatics of biomolecules contribute to their broad functionality. Developing a theoretical model to properly describe the electrostatics of proteins in water is the aim of the following sections.

In the last chapter, we introduced a linear response theory to account for the polarization effects of different media. The totally different composition of water-like and protein-like medium motivates to separate their dielectric behavior. Thus, this chapter is broken up into three sections:

Section 3.1 focuses on the description of the biomolecules we focus on in this work. This includes the definition of the molecule's surface, its charge distribution, and finally its dielectric response.

In Section 3.2 we discuss the solvent of biomolecules, namely water. Because of its special features, we go beyond the common treatment of macroscopic, constant dielectric response. Starting with the simplest situation, i.e., an infinite volume filled with pure water, we develop a nonlocal response theory, which incorporates the network character of the hydrogen bonds in the macroscopic response theory.

Merging the model for the biomolecule (Section 3.1) and the water model (Section 3.2) yields our description of the *biomolecular system*, which means a single biomolecule immersed in water. This is done in Section 3.3.

3.1 Model for biomolecules

A biomolecule is a molecule that is produced by a living organism. A special class of biomolecules are the so called *proteins*. Proteins are composed of amino acids [69], which are the building blocks of long polymer chains, see Fig. 3.1 for an illustration. With 2-10 amino acids such chains are called peptides, with 10-100 they are often called polypeptides, and longer chains are known as proteins. Proteins have many structural and enzymatic roles in organisms: they bind and recognize other biomolecules and this in turn causes new reactions. Proteins take part in building and repairing cell tissues, in water balancing or in nutrient transport, and induce muscle contractions, to mention just a few of their responsibilities.

Analyzing the functionality of proteins is an interesting field of research, since a deeper understanding can help to control and modify their functionality. Especially the highly specific binding or docking process of small biomolecules (ligands), which is a key point for drug development, is based on the electrostatic potential on the surface of the proteins and the electrostatic field emanated in the protein's surrounding [64].

In this section, we introduce the abstract description, i.e., the theoretical model, which is used in the remainder of this work to describe the electrostatics of a biomolecule. We start with the definition of the molecular charge distribution (Section 3.1.1), the description of the molecular surface (Section 3.1.2) and finally, we introduce the molecule's dielectric response (Section 3.1.3).

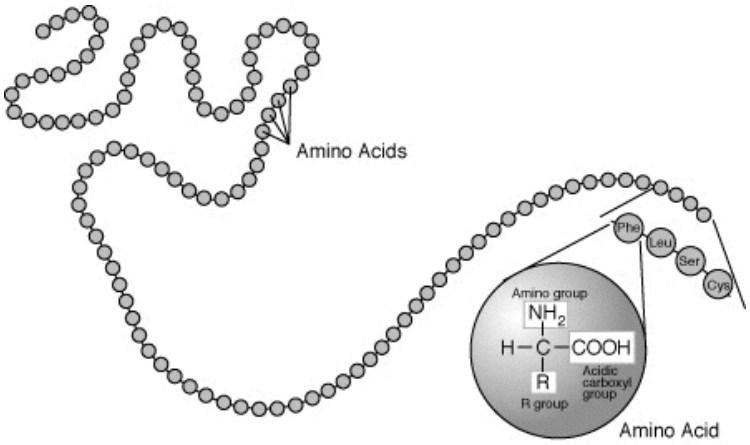


Fig. 3.1: Proteins are chains of amino acids.

3.1.1 Charge distribution of biomolecules

Atoms are complex objects, consisting of a number of protons, neutrons, and electrons. All of these subunits are quantum mechanical objects, which are defined rather by their probability distribution than by a fixed location in space. Molecules, which are composed of an arbitrary number of atoms, can be imagined to be even more complicated. This makes a description of proteins on the micromolecular scale as basis for a numerical model difficult.

Therefore, the atomic description is often reduced to merely two quantities, a radius r and a partial charge q: the value of the charge q is deduced from quantum mechanical calculations. It describes the excess charge of the atom in the molecular surrounding. The radius r describes a sphere, within which the probability to find the complete charge distribution of the atom equals a fixed percentage, e.g., 90% . In practice however, the radius is often chosen according to a force field to fit best to experimental solvation energies [78].

For a monoatomic molecule, this approximation yields a spherical shape. In the literature, we find different ways of distributing the charge q on or in the sphere [124]:

- The *Born sphere* defines the charge distribution to be a so called *point charge* in the center of the sphere. A point charge is physically expressed by an infinite charge density ρ at the point where the charge is located and zero everywhere else. The so called δ-distribution[1],

$$\begin{aligned} \delta(\boldsymbol{r}-\boldsymbol{r}') &= \begin{cases} \infty & \boldsymbol{r}=\boldsymbol{r}' \\ 0 & \boldsymbol{r}\neq\boldsymbol{r}' \end{cases} \\ \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\,\delta(\boldsymbol{r}-\boldsymbol{r}') &= 1\,, \end{aligned}$$

  allows for a mathematical description of a point charge,

$$\rho(\boldsymbol{r}) = q\delta(\boldsymbol{r}-\boldsymbol{r}')\,, \quad \text{in } \mathbb{R}^3$$

  which is located at $\boldsymbol{r}'$.

[1] We refer to [27, 28, 121] for an introduction to the theory of distributions.

- Another model for monoatomic ions is the so called *spherical shell*. Here, the charge q is homogeneously spread over the surface of the sphere of radius a. This results in the following constant surface charge density

$$\sigma(\boldsymbol{r}) = \frac{q}{4\pi a^2}\,\delta(r-a), \quad r = |\boldsymbol{r}|, \boldsymbol{r} \in \mathbb{R}^3.$$

Remark 3.1 These two models serve as analytically treatable test cases for the electrostatic models in Section 3.3.3. When describing realistic molecules, the Born spheres are taken to define its atoms, i.e., every atom i of the molecule is a sphere of radius a_i, where a fixed point charge q_i is localized at the center $\boldsymbol{r}_i$ of the sphere.

3.1.2 Surface description of biomolecules

The surface of the protein is an important quantity. The size and the shape of the surface influence a number of interactions contributing to energies such as the Van der Waals energy, entropic and excluded volume effects, the surface tension and also the electrostatic energy. In the introduction of this chapter, we motivated that the dielectric response of the protein in general differs from the dielectric response of the solvent. In an electrostatic context, this is why it is reasonable to define the surface of the molecule as the interface where the dielectric response changes, i.e., as the *dielectric boundary*.

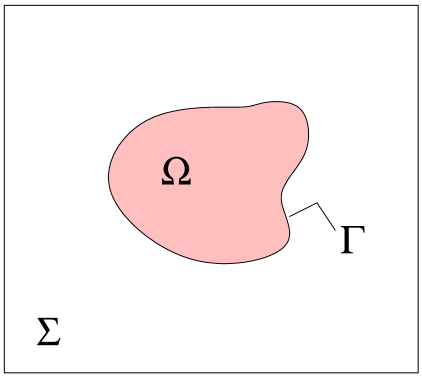


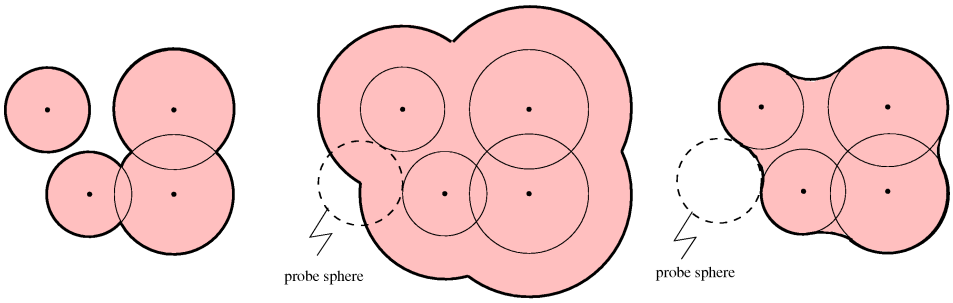
Fig. 3.2: Different domains of the molecular system.

We now introduce the different domains of the molecular system, which are crucial to describe the electrostatics of the system. The notations are illustrated in Fig. 3.2 and are used in this work to describe a molecule immersed in water: Ω denotes the inside of the molecule, where the dielectric response is given by the molecule. It is characterized and defined in Section 3.1.3. The molecule's surface is denoted by Γ. By definition, Γ separates Ω from the outer domain, which is characterized by the dielectric response of the solvent. Σ denotes the outer domain and the dielectric response in Σ is specified in Section 3.2.2.

The molecule's surface is not uniquely definable: starting with the question of the correct atom radius, which was introduced before, and ending with the question whether the surface is defined by the absence of the solvent or the presence of the molecular atoms.

There are three common surface definitions illustrated in Fig. 3.3 – all of them with different meaning and application. Consider a molecule consisting of a set of atoms. Each is represented as a Born sphere with given location and radius.

(a) The *Van der Waals Surface* (VdWS) is the hull of the union of all Born spheres the molecule consists of. The name of this surface is derived from the Van der Waals radii of the atoms.



(a) VdWS (b) SAS (c) SES

Fig. 3.3: Different surface descriptions of a molecule: the Van der Waals surface (VdWS), the solvent accessible surface (SAS) and the solvent excluded surface (SES).

(b) When a spherical solvent molecule is rolled over the VdWS, the *Solvent Accessible Surface* (SAS) is the surface traced out by the center of the solvent molecule during the rolling process.

(c) When a spherical solvent molecule is rolled over the VdWS, the *Solvent Excluded Surface* (SES) is the contact surface of the VdWS and the solvent molecule during the rolling process.

The SES of molecules in solution is often taken to define Γ as it separates the solvent molecules from the biomolecule. Additionally, for a numerical treatment this surface is appropriate because it is smooth almost everywhere. In the biomolecular studies presented in this work we use the SES, too. For all of the discussed surface representations the charge distribution lies completely in Ω and in the case of a monoatomic molecule the surface definitions, VdWS and SES, are equal.

3.1.3 Dielectric response of proteins

As discussed in Section 2.2.2, the electrostatic behavior in the interior of a medium is determined by its dielectric reaction and the fixed charges.

The inhomogeneous distribution of the atoms and the polar flexible groups at the protein's surface give rise to variations of the averaged polarization field $\boldsymbol{P}$ and therefore to variations of the biomolecule's dielectric response. In experimental and numerical studies, a stronger response has been observed at the surface of the protein: these variations are not specific interactions between the atoms, but are due to *local* changes in the composition of the protein [117]. In a polar solvent, polar side chains have the tendency to lie at the surface whereas nonpolar chains are buried deep inside the molecule. Such a local variance can be well described by the local, linear ansatz (2.24) proposed in Section 2.2.3 on p. 19, i.e., with the following, linear relation between the dielectric displacement field $\boldsymbol{D}$ and the electrostatic field $\boldsymbol{E}$:

$$\boldsymbol{D}(\boldsymbol{r}) = \varepsilon(\boldsymbol{r})\, \boldsymbol{E}(\boldsymbol{r})\,, \quad \text{in } \Omega$$

The dielectric function ε is measured in units of ε_0 and its spatial dependence accounts for local variation of the dielectric response. In the literature, one finds $\varepsilon(\boldsymbol{r})/\varepsilon_0$ to range between 2-10 for biomolecules [92]. Since we are primarily interested in the effects of water, we restrict ourselves to a constant macroscopic response of the biomolecule,

$$\varepsilon(\boldsymbol{r}) =: \varepsilon_{\text{macros,mol}} = 2\,\varepsilon_0\,.$$

3.1.4 Solvation energy

Based on the energy expressions derived in Section 2.3, we now focus on the energetic description of the biomolecular system. In thermodynamics, the energy of a system defines its macroscopic equilibrium state: this means that the energy determines whether the molecule is stable or if a reaction takes place. For instance the *free energy* difference of solvation measures the energy gain or energy loss in the case of changing the solvent environment of a molecule (e.g., from vacuum to water).

In Section 3.1.4.1 we derive the electrostatic part of the solvation energy. This quantity is of use in Section 3.3.3 for a comparison of theoretically predicted and experimental data. To define the complete free energy of the solvated molecule we have to consider energetic contributions which result from the molecule's shape and volume, i.e., nonpolar contributions. In Section 3.1.4.2 we characterize these nonpolar contributions and give an appropriate theoretical model to account for them.

3.1.4.1 Reaction field potential and electrostatic part of the solvation energy

The different dielectric responses of the molecule and its solvent are indicated by the different colors in Fig. 3.4. Furthermore, the figure illustrates the crucial idea of decomposing the electrostatic potential in two terms,

$$\phi(\boldsymbol{r}) = \phi_{reac}(\boldsymbol{r}) + \phi_{mol}(\boldsymbol{r})\,,$$

the so called *molecular field* ϕ_{mol} and the *reaction field* ϕ_{reac}.

The molecular field is created by the fixed charge distribution ρ embedded in the medium of molecular response ε_Ω. It fulfills the Laplace equation with source term ρ in $\mathbb{R}^3$

$$\triangle\phi_{mol}(\boldsymbol{r}) = -\frac{\rho(\boldsymbol{r})}{\varepsilon_\Omega}, \quad \boldsymbol{r} \in \mathbb{R}^3\,.$$

The solution of this equation has already been deduced in Section 2.1.1. It is proportional to the Newton potential of the Laplace operator $\triangle$,

$$\phi_{mol}(\boldsymbol{r}) = \frac{1}{\varepsilon_\Omega}(G_\triangle * \rho)(\boldsymbol{r}) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, \frac{\rho(\boldsymbol{r}')}{\varepsilon_\Omega} G_\triangle(\boldsymbol{r}-\boldsymbol{r}'), \quad \boldsymbol{r} \in \mathbb{R}^3\,. \tag{3.1}$$

The molecular field ϕ_{mol} is *independent* of the dielectric boundary and therefore yields the same contribution to ϕ in every setting of the same charge distribution.

The second part of ϕ, the reaction field ϕ_{reac}, measures the contribution due to the creation of

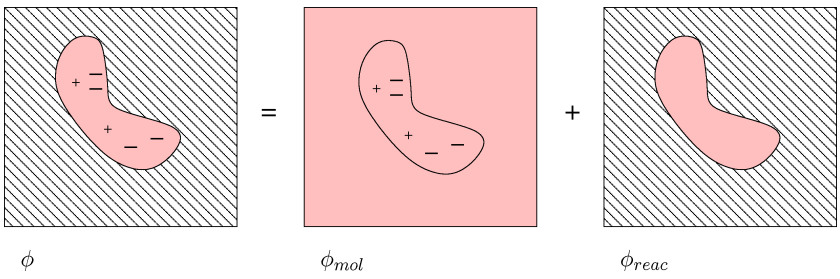


Fig. 3.4: The electrostatic potential is composed of the molecular potential and the reaction field potential. The pink region indicates the molecule with its dielectric response, whereas the white-lined surrounding represents the solvent.

the dielectric boundary[2]. Having this in mind, the *reaction field energy* can be interpreted as the energy gain when the molecule is transferred from its molecular surrounding (ϕ_{mol}) to the solvent ($\phi_{mol} + \phi_{reac}$).

With these notations, we deduce a general formula for the electrostatic contribution to the *solvation free energy* of molecules, $\triangle G_{elec}$. This energy is released or spent when transferring the molecule from one solvent, solvent I, into another solvent, solvent II. Thus, the solvation energy is given by the difference[3] between the field energies

$$\triangle G_{elec} = E_{field,II} - E_{field,I}\,.$$

In both solvents, the field energies, $W_{field,I}$ and $W_{field,II}$ are given by Eq. (2.29)

$$W_{field} = \int_{\mathbb{R}^3} d\boldsymbol{r} \int_0^{\boldsymbol{D}} \delta \boldsymbol{D}(\boldsymbol{r}) \cdot \boldsymbol{E}(\boldsymbol{r})\,,$$

where $\boldsymbol{E}$ and $\boldsymbol{D}$ are the solutions of the material equations for a given dielectric operator ε. These field energies originate from the same charge distribution. The only fact that differs is the dielectric boundary, which contributes to the energy by ϕ_{reac}. When assuming the charge distribution to completely lie in Ω and further assuming a local dielectric function, the electrostatic contribution to the solvation energy is given by

$$\begin{aligned} \triangle G_{elec} &= \frac{1}{2}\int_\Omega \rho(\phi_{II} - \phi_I) \\ &= \frac{1}{2}\int_\Omega \rho\left((\phi_{mol} + \phi_{reac,\,II}) - (\phi_{mol} + \phi_{reac,\,I})\right) \\ &= \frac{1}{2}\int_\Omega \rho\left(\phi_{reac,\,II} - \phi_{reac,\,I}\right)\,. \end{aligned} \tag{3.2}$$

This is an important result, since it tells us that only the reaction field energies contribute to the solvation energy. The molecular field energy comprises the infinite self energy terms, which we discussed in Section 2.3. These terms cancel out in Eq. (3.2) and therefore we end up with physically interpretable energies.

Important remark 3.1 We learned in Section 2.2.2 that a change in the dielectric response of the biomolecular system, results in a jump in the polarization, which in turn can be interpreted as induced surface charge. Having this in mind, the reaction field *is* the electrostatic field, which is created by these surface charges. In the literature, many sophisticated and elegant methods are discussed to calculate the reaction field potential and the reaction field energies or appropriate approximations directly from a given charge setting, see [43, 71, 101] and the overview on binding energy calculations in Appendix 10.4.

3.1.4.2 Nonpolar energy contributions

The previous discussion has been about the electrostatic contribution to the solvation free energy. This is an important, but not the only contribution to the energy of a thermodynamic system. In order to complete the discussion on the solvation free energy, we now give a brief overview of the missing contributions and refer to supplementary literature.

[2]Mathematical view: in Section 2.1.2 we introduced a correction term for boundary value problems to guarantee the correct boundary values, ϕ_{reac} *is* this correction term (single/double layer).

[3]Here, $\triangle$ implies a difference.

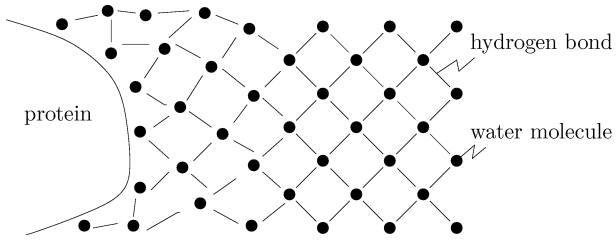


Fig. 3.5: Creation of the biomolecule's cavity destroys hydrogen bonds and leads to a loss in entropy.

In addition to the electrostatic free energy, the solvation free energy is influenced by nonpolar effects. These entropic as well as enthalpic contributions are relevant in the first shells around the solvated molecules. The following decomposition is for instance proposed in [13]:

$$\Delta G_{solv} = \Delta G_{elec} + \Delta G_{vdW} + \Delta G_{H\text{-}bond} + \Delta G_{cav} \tag{3.3}$$

These are the electrostatic energy, which we discussed before, the van-der-Waals energy between the solvated molecule and the neighboring solvent molecules, and the H-bond contributions of electronegative atoms on the surface. In contrast to these enthalpic contributions, the cavity term ΔG_{cav} is of entropic nature. It describes the mechanical work needed for creating a cavity in the solvent that contains the solvated molecule. One can imagine that the cavity term is high for water, whose network is partially destroyed or reoriented during and after the solvation process. An exact estimation of this effect is difficult, because the magnitude highly depends on the local composition of the molecule, which is solvated. Fig. 3.5 depicts the reorganization of the hydrogen bonds when approaching an interface which is not capable of building hydrogen bonds. This is for example the case when considering an overall uncharged molecule with mainly nonpolar side chains. Then, the *hydrophobic* effect. i.e., the energy needed for the reorganization of the water molecules, can dominate over the electrostatic contribution [119].

In the literature, various approaches are proposed for an incorporation of the nonpolar contributions to the solvation energy of biomolecules, see for example [67, 135] and citations given therein. From its magnitude, the cavity term is the most important nonpolar contribution in Eq. (3.3) at least for proteins. In order to avoid complicated but insufficiently reliable microscopic calculations involving new adjustable parameters, one often uses a quasi-macroscopic approach, such as the so called *empirical Uhlig formula*, where the solubility of nonpolar solutes is related to the molecular surface area of the solute and the interfacial tension [37, 132]. Another approach is given in [130], where the change in solvent pressure is calculated when building up the cavity. These so called *scaled particle theories* indicate that cavity creation work should depend on both, the solvent accessible volume and the solvent accessible surface of the molecule [62, 135].

As explained above, the reorganization and the breaking of the hydrogen bond network in the first solvation shell of the molecule plays an important role in the entropic contributions to the solvation energies of biomolecules. In the present work, we develop a novel formulation of the dielectric response of water, which takes into account the correlations of water molecules due to their hydrogen bond network. This model, in principle, allows for a specification of the correlation effect at the interface and therefore can in parts incorporate the cavity energy terms, see Remark 4.4 on p. 63. The calculation and the detailed investigation of these energy contributions in the framework of nonlocal electrostatics has still to be studied and constitutes a focus of future work.

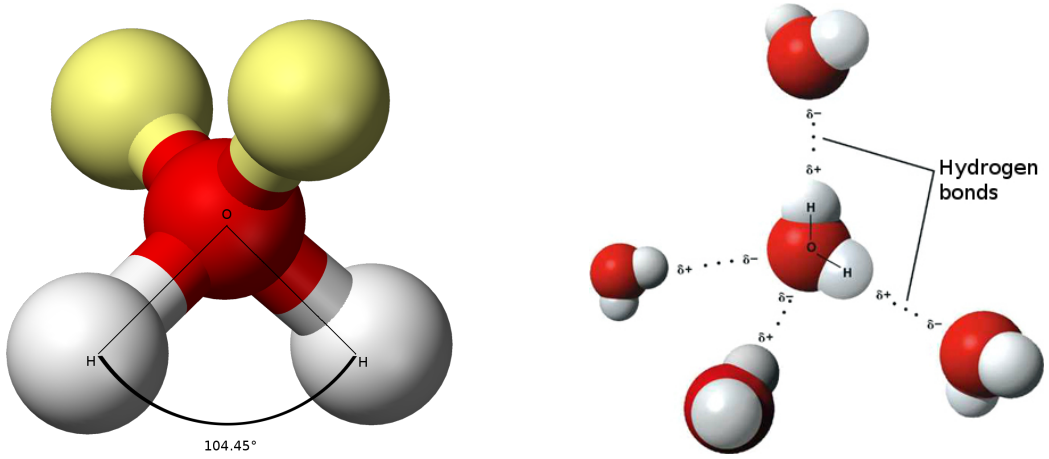


Fig. 3.6: left: tetrahedral structure of the water molecule, the two lone electron pairs are drawn in yellow; right: 3D water network.

3.2 Water - the main solvent of biomolecules

All the major components in cells, such as proteins, DNA or polysaccharides, are solvated in water. This is the reason why water is of immense interest in biological problems. It is important to understand the characteristics of water to map biochemical systems correctly into a theoretical framework. In Section 3.2.1 we summarize the main features of this multi-faceted solvent and in Section 3.2.2 we derive a theoretical model to describe its dielectric response.

3.2.1 Characteristics of water

A single molecule of water has two hydrogen atoms, which are covalently bound to a single oxygen atom. The latter possesses two lone electron pairs. Because of the quantum mechanical distribution of the electron pairs in the chemical bonds, the H-O-H bond angle is a distorted tetrahedral angle of about 104.45° causing a three-dimensional molecular structure. This is illustrated in Fig. 3.6 on the left. The total charge of water is zero. However, as the water molecule is non-linear and as the oxygen atom has a higher electronegativity than the hydrogen atoms, the former carries a slightly negative excess charge whereas the latter are slightly positive. As a result, water is a polar molecule with a permanent electrical dipole moment.

In addition to the dipolar nature, which causes the water molecule to be a proton donor on the two hydrogen ends, the water molecule happens to be a proton acceptor, as well. This also originates from two lone electron pairs present in the outer shell of the oxygen. Between these two functional groups, the proton donor and acceptor, the water forms so called *hydrogen bonds*. As it is shown on the right of Fig. 3.6, a water molecule can form up to four hydrogen bonds with surrounding water molecules.

Without an external electrostatic field the water molecules try to maximize their hydrogen bonds to neighboring molecules. Because of thermal motion and local deviations it is reasonable to introduce a *correlation length* λ, which defines the greatest distance at which two water molecules still influence each other by the hydrogen network, see Fig. 3.7.

If we consider the water molecules to be exposed to an external electrostatic field, there are two competing effects: on the one hand, the water molecules try to keep the hydrogen bonds with the surrounding molecules. On the other hand, the water molecules individually try to align with the external field to minimize the potential energy of the molecular dipoles.

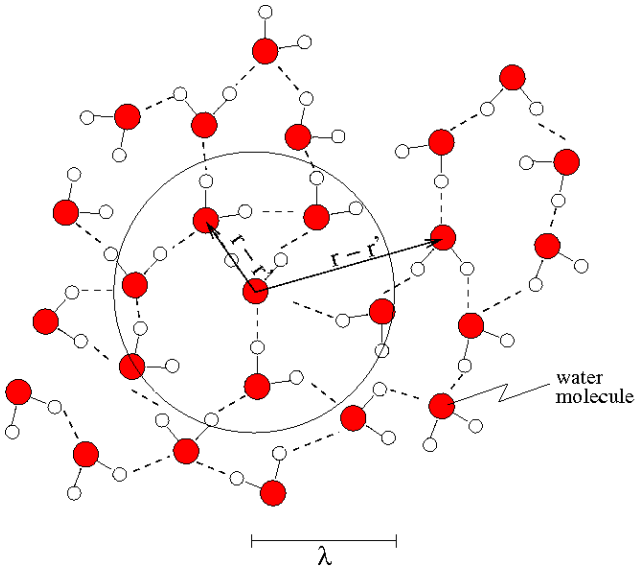


Fig. 3.7: Correlation length of the water network.

In numerical simulations there are two different tools to represent water. The first is to model the water molecules explicitly in order to capture the discussed effects directly. Popular water models used in molecular dynamic simulations are, for instance, the TIPNP models and the SPC\E, see [20, 47, 137] for a review and [77] for a general overview. They differ in handling the structure as rigid or flexible, in the number of possible interactions, and in whether the model includes polarization effects. The computational costs are high and increase with the number of water molecules taken into account. An alternative to the explicit water models is to use an *implicit solvation model*, also known as a continuum model, within which water is treated as a dielectric medium. Its reaction to the presence of an external electrostatic field $\boldsymbol{E}$ causes a macroscopic, effective polarization $\boldsymbol{P}$.

3.2.2 The dielectric response of water

In this section we focus on an implicit solvation model to describe the peculiar dielectric response which originates from the permanent dipoles and the hydrogen bond network of water. First, we assume the water to capture the whole space. In this case, the dielectric response has to be invariant under translation or rotation and this simplifies the modeling process. Later, these assumptions turn out to be approximations for the molecular system, where the immersed molecule apparently is a symmetry-breaking boundary.

3.2.2.1 The nonlocal dielectric operator

For a first approximation of the dielectric response of water, Eqs. (2.24)(a-c),

$$\begin{aligned} \boldsymbol{P}(\boldsymbol{r}) &= \chi(\boldsymbol{r})\,\boldsymbol{D}(\boldsymbol{r}), && \boldsymbol{r} \in \mathbb{R}^3 \\ \boldsymbol{P}(\boldsymbol{r}) &= (\varepsilon(\boldsymbol{r}) - \varepsilon_0)\,\boldsymbol{E}(\boldsymbol{r}), && \boldsymbol{r} \in \mathbb{R}^3 \\ \boldsymbol{D}(\boldsymbol{r}) &= \varepsilon(\boldsymbol{r})\,\boldsymbol{E}(\boldsymbol{r}), && \boldsymbol{r} \in \mathbb{R}^3 \\ \text{with} \quad \chi(\boldsymbol{r}) &= (1 - \varepsilon^{-1})(\boldsymbol{r}), && \boldsymbol{r} \in \mathbb{R}^3\,. \end{aligned}$$

are often used. The constant, macroscopic response of water at room temperature is $\varepsilon_{\text{macros,wat}} = 78.5\,\varepsilon_0$. Indeed, in the majority of cases, numerical solvers applied in biomolecular studies use this simple ansatz [8, 45, 129].

However, as we already anticipated in Section 2.2.3, the hydrogen bond network causes non-trivial correlations of water molecules and their length scale exceeds the atomic scale. To account for these correlations, we have to start with the linear, nonlocal response theory given by Eqs. (2.25)(a-c) [74],

$$P_i(t,\boldsymbol{r}) = \int_{-\infty}^{t}\int_{\mathbb{R}^3} \left(\varepsilon_{ij}(t-\tau;\boldsymbol{r},\boldsymbol{r}') - \varepsilon_0\delta_{ij}\right) E_j(\tau,\boldsymbol{r}')d\boldsymbol{r}'d\tau, \qquad \boldsymbol{r}\in\mathbb{R}^3$$

$$P_i(t,\boldsymbol{r}) = \int_{-\infty}^{t}\int_{\mathbb{R}^3} \chi_{ij}(t-\tau;\boldsymbol{r},\boldsymbol{r}')D_j(\tau,\boldsymbol{r}')d\boldsymbol{r}'d\tau, \qquad \boldsymbol{r}\in\mathbb{R}^3$$

$$D_i(t,\boldsymbol{r}) = \int_{-\infty}^{t}\int_{\mathbb{R}^3} \varepsilon_{ij}(t-\tau;\boldsymbol{r},\boldsymbol{r}')E_j(\tau,\boldsymbol{r}')d\boldsymbol{r}'d\tau, \qquad \boldsymbol{r}\in\mathbb{R}^3\,.$$

In general, the tensor functions, ε_{ij} and χ_{ij} can be very complicated. However, assuming the water to fill the whole space, there is no reason to assume that a translation or a rotation macroscopically changes the basic features of a hydrogen bond. We can simplify Eqs. (2.25)(a-c) in the following way:

- We do not assume a time dependency of the dielectric reaction, since we focus on equilibrated, thermodynamically stable systems. This is why we will neglect the variations on t and τ:

$$D_i(\boldsymbol{r}) = \int_{\mathbb{R}^3} \varepsilon_{ij}(\boldsymbol{r},\boldsymbol{r}')E_j(\boldsymbol{r}')d\boldsymbol{r}', \quad \boldsymbol{r}\in\mathbb{R}^3$$

- We assume that the correlations of the water molecules do not depend on the position in space, i.e., the dielectric function is invariant under an arbitrary translation and therefore homogeneous:

$$D_i(\boldsymbol{r}) = \int_{\mathbb{R}^3} \varepsilon_{ij}(\boldsymbol{r}-\boldsymbol{r}')E_j(\boldsymbol{r}')d\boldsymbol{r}', \quad \boldsymbol{r}\in\mathbb{R}^3$$

- We assume that the nonlocal correlations of the water molecules at position $\boldsymbol{r}$ do not depend on the direction which is specified by $(\boldsymbol{r}-\boldsymbol{r}')$, but that ε_{ij} is invariant under an arbitrary rotation and therefore isotropic. This means that only the distance $|\boldsymbol{r}-\boldsymbol{r}'|$ is crucial for the dielectric response [4]:

$$D_i(\boldsymbol{r}) = \int_{\mathbb{R}^3} \varepsilon_{ij}(|\boldsymbol{r}-\boldsymbol{r}'|)E_j(\boldsymbol{r}')d\boldsymbol{r}', \quad \boldsymbol{r}\in\mathbb{R}^3 \tag{3.4}$$

All in all, we assume with Eq. (3.4) an *isotropic, linear, nonlocal* dielectric function for water capturing the whole space.

[4] Except for dependencies of the form $\delta(\boldsymbol{r}-\boldsymbol{r}')$. In this case the integral description is reduced to a constant, local response function, which is, by definition, homogeneous and isotropic.

3.2.2.2 Constraints on the nonlocal dielectric function

A closer look at Eq. (3.4) reveals that the action of the dielectric operator is a convolution:

$$\boldsymbol{D}_i(\boldsymbol{r}) = (\varepsilon_{ij} * \boldsymbol{E}_j)(\boldsymbol{r}) = \int_{\mathbb{R}^3} \varepsilon_{ij}(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{E}_j(\boldsymbol{r}')d\boldsymbol{r}', \quad \boldsymbol{r} \in \mathbb{R}^3$$

This suggests that the discussion about the features of the dielectric operator is done best in Fourier space, as in Fourier space the convolution turns into a multiplication. With the notation $\hat{\mathcal{Q}}$ for the Fourier transformation $\mathcal{F}(\mathcal{Q})$ of a quantity $\mathcal{Q}(\boldsymbol{r})$, the dielectric field $\boldsymbol{D}$ in Fourier space reads:

$$\hat{\boldsymbol{D}}_i(\boldsymbol{k}) = \hat{\varepsilon}_{ij}(\boldsymbol{k})\,\hat{\boldsymbol{E}}_j(\boldsymbol{k}), \quad \boldsymbol{k} \in \mathbb{R}^3 \tag{3.5}$$

The dispersion relation is reduced to a *local* dependency on the wave vector $\boldsymbol{k}$. Here, the wave length $\frac{2\pi}{k}$ determines the critical distance where the field still has essential variations. Having this in mind, the dispersion relation expresses the dependency of the macroscopic material property on the field variations in space.

Before we state an appropriate model for the dielectric function $\hat{\varepsilon}$ of water, we discuss some features that $\hat{\varepsilon}$ has to fulfill:

Reduction to the longitudinal component: In presence of a $\boldsymbol{k}$-dependency as given in Eq. (3.5), the dielectric function, $\hat{\varepsilon}_{ij}$ is a tensor of second order and does not reduce to a scalar. For an isotropic medium, one can split the tensor functions into their longitudinal ($\hat{\varepsilon}_{\parallel}$) and transversal components ($\hat{\varepsilon}_{\perp}$), which only depend on the absolute value of the wave vector, $k = |\boldsymbol{k}|$ [74]:

$$\hat{\varepsilon}_{ij}(\boldsymbol{k}) = \hat{\varepsilon}_{\parallel}(k)\frac{k_i\,k_j}{k^2} + \hat{\varepsilon}_{\perp}(k)\,(\delta_{ij} - \frac{k_i\,k_j}{k^2}) \tag{3.6}$$

Assuming a fixed charge density ρ in a homogeneous, isotropic, nonlocal medium, the material equations given by Theorem 2.2.1 read

$$\begin{aligned}
& & -\tfrac{\partial}{\partial r_i}\left(\int d\boldsymbol{r}'\,\varepsilon_{ij}(\boldsymbol{r}-\boldsymbol{r}')\tfrac{\partial}{\partial r'_j}\phi(\boldsymbol{r}')\right) &= \rho(\boldsymbol{r}) \\
\Rightarrow & & k_i\left(\hat{\varepsilon}_{ij}(\boldsymbol{k})k_j\hat{\phi}(\boldsymbol{k})\right) &= \hat{\rho}(\boldsymbol{k}) \\
\text{with (3.6)} & \sum_{i,j}\left(\hat{\varepsilon}_{\parallel}(k)\tfrac{k_i^2k_j^2}{k^2} + \hat{\varepsilon}_{\perp}(k)(k_ik_j\delta_{ij} - \tfrac{k_i^2k_j^2}{k^2})\right)\hat{\phi}(\boldsymbol{k}) &= \hat{\rho}(\boldsymbol{k}) \\
& & \left(\hat{\varepsilon}_{\parallel}(k)\tfrac{k^2k^2}{k^2} + \hat{\varepsilon}_{\perp}(k)(k^2 - \tfrac{k^2k^2}{k^2})\right)\hat{\phi}(\boldsymbol{k}) &= \hat{\rho}(\boldsymbol{k}) \\
\Leftrightarrow & & \hat{\phi}(\boldsymbol{k}) = \hat{\rho}(\boldsymbol{k})/\left[\hat{\varepsilon}_{\parallel}(k)\,k^2\right]\,. &
\end{aligned}$$

Thus, for a linear, but nonlocal, response theory it suffices to find an appropriate model for the longitudinal component $\hat{\varepsilon}_{\parallel}$ of the dielectric tensor, which depends on k. In the following, we skip the subscript $\parallel$ and refer to it as $\hat{\varepsilon}$.

In summary, we find that for electrostatic problems the tensor in Eq. (3.5) is a scalar function

$$\hat{\boldsymbol{D}}(\boldsymbol{k}) = \hat{\varepsilon}(k)\,\hat{\boldsymbol{E}}(\boldsymbol{k}), \quad k = |\boldsymbol{k}|, \boldsymbol{k} \in \mathbb{R}^3\,. \tag{3.7}$$

Thermodynamic limits: Further insight into the behavior of $\hat{\varepsilon}$ is obtained from thermodynamic requirements: in general, both, $\hat{\chi}(\boldsymbol{k},\omega)$ and $\hat{\varepsilon}(\boldsymbol{k},\omega)$, are complex-valued functions. As $\boldsymbol{D}$ and $\boldsymbol{P}$ are the independent thermodynamic variables of the system, χ is a *general susceptibility* which has to fulfill Kramers-Kronig relation [6, 75, 86]. The imaginary part of $\hat{\chi}$ describes

the loss in field energy by excitations of inner degrees of freedom, whereas the real part of $\hat{\chi}$ describes the effective polarization of the medium as a function of the external field $\boldsymbol{D}$. In the stationary case, i.e., setting $\omega = 0$, we can assume $\hat{\chi}$ and $\hat{\varepsilon}$ to be real quantities, as the loss in energy is mainly caused by the frequency dependent excitation of internal degrees of freedom. For the static dielectric function $\hat{\varepsilon}(k) = \Re(\hat{\varepsilon}(k, \omega = 0))$, the Kramers-Kronig relation then reads

$$\hat{\chi}(k) = \Re(\hat{\chi}(k)) = 1 - \frac{1}{\sqrt{2\pi}^3 \hat{\varepsilon}(k)} = \frac{2}{\pi} \int_0^\infty \frac{d\omega}{\omega} \frac{\Im(\hat{\varepsilon}(k,\omega)}{|\hat{\varepsilon}(k,\omega)|^2} \,.$$

The energy balance in equilibrium systems requires $\Im(\hat{\varepsilon}(k, \omega)) > 0$, meaning that the medium cannot create energy and dissipate it into the system. Therefore, the dielectric function is subject to the restriction:

$$\hat{\chi}(k) > 0 \qquad \Rightarrow \qquad \varepsilon(k) > \frac{1}{\sqrt{2\pi}^3} \quad \text{or} \quad \varepsilon(k) < 0$$

As a general susceptibility, χ is subject to the fluctuation-dissipation theorem and therefore is connected to the polarization fluctuation, $\langle P_i(\boldsymbol{r}, t)\, P_k(\boldsymbol{0}, 0)\rangle$, which causes the medium to react to the external field [75]. A quantity that captures the effects of the polarization fluctuation as a function of wave number and frequency is the so called form factor of the polarization fluctuation

$$\begin{aligned} S(\boldsymbol{k}, \omega) &= \frac{1}{2\pi^3} \frac{k_i\, k_j}{k^2} \int_V e^{-i(\boldsymbol{k}\boldsymbol{r} - \omega t)} \langle P_i(\boldsymbol{r}, t)\, P_k(\boldsymbol{0}, 0)\rangle \,, \\ \hat{\chi}(k) &= 1 - \frac{1}{\sqrt{2\pi^3}\hat{\varepsilon}(k)} = 4/\hbar \int_0^\infty \frac{d\omega}{\omega} [1 - e^{-\beta\hbar\omega}] S(\boldsymbol{k}, \omega) \,. \end{aligned}$$

The latter equation is a very important relation, as it turns out to be a tool for the determination of $\hat{\varepsilon}(k)$ on the basis of $S(k)$. Data for $S(k)$ is available from isotopic substitution method for the analysis of diffraction experiments [118] and from computer simulations [124].

Local limit constraints: For any dielectric medium it holds that

$$\lim_{k \to 0} \hat{\varepsilon}(k) = const = \frac{1}{\sqrt{2\pi}^3}\, \varepsilon_{\text{macros}} \,, \tag{3.8}$$

where $\varepsilon_{\text{macros}}$ is the macroscopic, constant response of the medium. This means that in the case of vanishing field variations, i.e., when applying a constant electrostatic field, the solvent molecules would homogeneously shield the field. The magnitude of the shielding corresponds to the macroscopic value of the orientational polarization: the water network is arranged in a way that the permanent dipoles of the water molecule effectively rotate in the external field to minimize their energy. The macroscopic response of water at room temperature is commonly described by

$$\varepsilon_{\text{macros,wat}} = 78.5\, \varepsilon_0 \,.$$

In contrast, for high, spatial variations of the electric field, the permanent dipoles can no longer align with the field, as the spatial variations are smaller than the finite size of the permanent dipole. Then, the dielectric function reaches the limit

$$\lim_{k \to \infty} \hat{\varepsilon}(k) = const = \frac{1}{\sqrt{2\pi}^3}\, \varepsilon_\infty \,, \tag{3.9}$$

where ε_∞ characterizes the polarization field $\boldsymbol{P}$ in terms of electronic internal degrees of freedom such as variations in the overlap of orbitals and of the electron probability. In the case of water, the high-k limit is taken to be equal to the square of the refractive index. It can then be calculated from the Clausius-Mossotti equation [19] and reveals [50, 124, 137]:

$$\varepsilon_\infty = 1.8\,\varepsilon_0$$

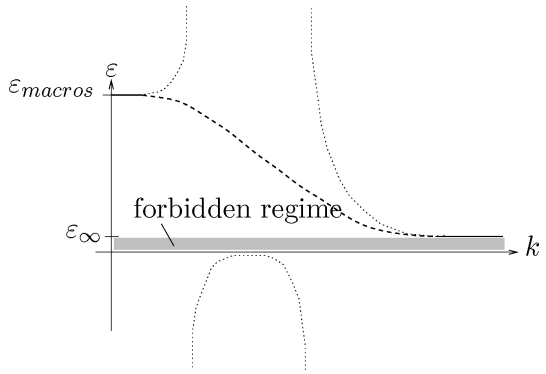


Fig. 3.8: Two possible shapes of $\hat{\varepsilon}(k)$.

A model for the dielectric response of water has to fulfill all these restrictions in Fourier space. Two possible functional forms of $\hat{\varepsilon}$ are shown in Fig. 3.8. The dotted curve has two divergence points and a region of negative dielectric response. Negative values for the dielectric function were reported first for molten salt systems, where the ions form a highly correlated system, which *overscreens* the fixed charge. The ions form alternating shells with negative and positive ions around a central ion. This results in local electrostatic fields opposite to the external driver. For dielectric, polar media a few theoretical studies and numerical simulations also predict a k-region of negative response, see [86, 102, 124] and references therein.

Although the approaches that admit a negative dielectric function are in good agreement when comparing the static form factor of the polarization fluctuations, they fail in correctly describing experimentally measurable, macroscopic quantities such as the hydration energy or the solvation energies of small ions [102, 124]. This can be due to the different scales on which the physical effects are compared: polarization fluctuations and the resulting form factor live on the nano-scale, whereas energies belong to the macroscopic quantities of a system. In contrast to molten salt or ideal dipolar fluids, real polar media have an internal molecular structure, which reacts sensitively to developed hydration shells. Moreover, the electrostatic field of a static charge distribution is not completely screened by water molecules as can be seen by considering the macroscopic dielectric constant of water, which coincides with a strong, but never complete, screening.

Therefore, we restrict ourselves to

$$\varepsilon(k) > \frac{1}{\sqrt{2\pi}^3}\,,$$

i.e., the polarization effect is always weaker than the effect of the fixed charges such as the dashed $\hat{\varepsilon}$ function in Fig. 3.8. This model is now discussed in detail.

3.2.3 The Lorentzian model of water

A dielectric function $\hat{\varepsilon}$, which fulfills all the deduced constraints, is proposed in [33,35] and is given by Eq. (3.10).

$$\frac{1}{\hat{\varepsilon}(k)} = \sqrt{2\pi}^3 \left(\frac{1}{\varepsilon_\infty} - (\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_{\text{macros}}}) \frac{1}{1+\lambda^2\, k^2} \right), \qquad k \in \mathbb{R} \tag{3.10}$$

$$\Leftrightarrow \quad \hat{\varepsilon}(k) = \frac{1}{\sqrt{2\pi}^3} \left(\varepsilon_\infty + (\varepsilon_{\text{macros}} - \varepsilon_\infty) \frac{1}{1+\frac{\varepsilon_{\text{macros}}}{\varepsilon_\infty}\lambda^2\, k^2} \right), \qquad k \in \mathbb{R} \tag{3.11}$$

The dashed line in Fig. 3.8 sketches the principal shape of $\hat{\varepsilon}(k)$: it is the sigmoidal connection between the two frequency limits: a gradual decrease in screening of the electric field from the high macroscopic response (large distance scales, small k) to the low dielectric constant (small distance scales, large k), when orientational degrees of freedom in the solvent do not respond to the electric field anymore. This model for the dielectric operator is called the *Lorentzian* model, where the name is coined from the shape of Eq. (3.11) in real space.

In Eq. (3.10) the correlation length λ arises. This quantity has already been motivated in Section 3.2.1: λ determines the length scale on which the local hydrogen-bonded water correlations decrease. The basis of this model is to assume an exponential decay of the spatial polarization fluctuation due to the low-frequency part of the Debye spectrum [33]. The exponential decay is characterized by λ. Introducing such a correlation length goes qualitatively beyond the macroscopic effects and it is a first step to incorporate non-trivial polarization correlations based on the collective network character of water.

Since the polarization fluctuations lead to a deviation from the macroscopic response, we expect for $\lambda \to 0$ to regain the local response given in Eq. (3.8). Then, the water molecules effectively align with the external electrostatic field as it is the case for $k \to 0$. Indeed, assuming the Lorentzian model, the limiting process yields

$$\lim_{\lambda\to 0} \hat{\varepsilon}(k) = \frac{1}{\sqrt{2\pi}^3}\, \varepsilon_{\text{macros}}\,. \tag{3.12}$$

In contrast, the limiting process, $\lambda \to \infty$, implying a very large correlation length compared to the variations of the electrostatic field, reveals the same electronic response as given in Eq. (3.9) for the high-k limit: the water molecules build up a stable water network and their reaction due to the external field is exclusively characterized by their electronic polarization,

$$\lim_{\lambda\to \infty} \hat{\varepsilon}(k) = \frac{1}{\sqrt{2\pi}^3}\, \varepsilon_{\infty}\,.$$

From Eq. (3.11) we see that it is reasonable to introduce a new quantity, namely the scaled reciprocal of the correlation length

$$\kappa := \frac{1}{\lambda} \sqrt{\frac{\varepsilon_\infty}{\varepsilon_{\text{macros}}}}\,. \tag{3.13}$$

To analyze the dielectric operator $\varepsilon(\boldsymbol{r}-\boldsymbol{r}')$ in real space, we transform $\hat{\varepsilon}(k)$ back. As a detailed derivation of the transformation is given in [52], we only state the result here:

$$\hat{\varepsilon}(k) = \frac{1}{\sqrt{2\pi}^3} \left(\varepsilon_\infty + (\varepsilon_{\text{macros}} - \varepsilon_\infty) \frac{1}{1+k^2/\kappa^2} \right), \qquad k = |\boldsymbol{k}|, \boldsymbol{k} \in \mathbb{R}^3$$

$$\varepsilon(\boldsymbol{R}) = \varepsilon_\infty \delta(\boldsymbol{R}) + (\varepsilon_{\text{macros}} - \varepsilon_\infty) \frac{\kappa^2}{4\pi} \frac{e^{-R\kappa}}{R}, \qquad R = |\boldsymbol{R}|, \boldsymbol{R} \in \mathbb{R}^3$$

Identifying $\boldsymbol{R}$ with the difference of arbitrary points in the dielectric medium,

$$\boldsymbol{R} = \boldsymbol{r} - \boldsymbol{r}' \,,$$

we end up with the nonlocal dielectric operator and the integral description of the dielectric permittivity, respectively:

$$\varepsilon(\boldsymbol{r}-\boldsymbol{r}') = \varepsilon_\infty \delta(\boldsymbol{r}-\boldsymbol{r}') + (\varepsilon_{\text{macros}} - \varepsilon_\infty)\frac{\kappa^2}{4\pi}\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}, \qquad \boldsymbol{r},\boldsymbol{r}' \in \mathbb{R}^3 \tag{3.14a}$$

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{r}) &= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' \varepsilon(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{E}(\boldsymbol{r}) \\ &= \varepsilon_\infty \boldsymbol{E}(\boldsymbol{r}) + (\varepsilon_{\text{macros}} - \varepsilon_\infty) \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' \frac{\kappa^2}{4\pi}\frac{e^{-\kappa\,|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\boldsymbol{E}(\boldsymbol{r}'), \qquad \boldsymbol{r} \in \mathbb{R}^3\,. \end{aligned} \tag{3.14b}$$

Since $\hat{\varepsilon}$ fulfills the physical constraints given in Eqs. (3.8), (3.9), and (3.12) the corresponding ε fulfills them as well. For instance, we regain the local limit for $\kappa \to \infty$ in $\mathbb{R}^3$, because

$$\lim_{\kappa\to\infty} \frac{\kappa^2}{4\pi}\left(\frac{e^{-\kappa\,|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\right) = \begin{cases} 0 & \text{for} \quad \boldsymbol{r} \neq \boldsymbol{r}' \\ \infty & \text{for} \quad \boldsymbol{r} = \boldsymbol{r}' \end{cases} \quad \text{for all } \kappa > 0 \text{ and } \boldsymbol{r}' \in \mathbb{R}^3 \tag{3.15a}$$

$$\int\limits_{\mathbb{R}^3} d\boldsymbol{r}' \frac{\kappa^2}{4\pi}\frac{e^{-\kappa\,|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} = 1, \qquad \text{for all } \kappa > 0\,, \tag{3.15b}$$

where the last identity can be proven using mathematical software tool such as Maple or Mathematica. As Eqs. (3.15)(a,b) define a δ-distribution in $\mathbb{R}^3$, we find

$$\begin{aligned} \lim_{\kappa\to\infty} \varepsilon(\boldsymbol{r}-\boldsymbol{r}') &= \varepsilon_\infty\, \delta(\boldsymbol{r}-\boldsymbol{r}') + (\varepsilon_{\text{macros}} - \varepsilon_\infty) \lim_{\kappa\to\infty} \frac{\kappa^2}{4\pi}\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} \\ &= \varepsilon_\infty\, \delta(\boldsymbol{r}-\boldsymbol{r}') + (\varepsilon_{\text{macros}} - \varepsilon_\infty)\delta(\boldsymbol{r}-\boldsymbol{r}') \\ &= \varepsilon_{\text{macros}}\, \delta(\boldsymbol{r}-\boldsymbol{r}')\,, \qquad \boldsymbol{r},\boldsymbol{r}' \in \mathbb{R}^3\,. \end{aligned}$$

The Lorentzian dielectric operator is the simplest approach which fulfills the thermodynamic limits of a nonlocal response theory, i.e., a sigmoidal connection between the limits of low and high wave number. Despite of its simplicity, which cannot withstand certain criticism [6], first applications on simple and non-trivial geometries demonstrate the success of this ansatz (see Section 3.3.3).

After a brief recapitulation of previous work, we focus on a new, precise definition of the dielectric response in biomolecular systems and a deeper understanding of the Lorentzian model of nonlocal electrostatics.

Historical review

- In 1970, the pioneer work on nonlocal electrostatics was done by Kornyshev et al. [72] and Vorotyntsev [134], who transferred the dielectric response theory of semiconductor research to a model for the nonlocal effect of water. This work was inspired by Inkson's simple model for the semiconductor dielectric response [59].
- Kornyshev et al. solved the integro-differential equations for the planar surface and for spher-

ical symmetry. This could be achieved, as in these cases the three-dimensional problem can be reduced to one degree of freedom and therefore is analytically solvable, see Section 3.3.3.

- In 2005, A. Hildebrandt found a more elegant and simpler method to solve the integro-differential system: he identified the integral kernel as the fundamental solution of a special, well-known differential operator [52]. This leads to the possibility to transform the integro-differential system to a purely differential system.

The present work focuses on the differential approach proposed by A. Hildebrandt. We embed the ideas given in [52] into a mathematical framework, which allows for the first time to modify and thoroughly interpret the established nonlocal equations: the general formulation, which we will derive, gives the possibility to handle the biomolecule as a "disruptive" element within the water network. This is relevant, as on the one hand the water network could be weakened or destroyed because of the physical boundary of the molecule. On the other hand, strong interactions might appear between individual water molecules and surface exposed amino acids in the form of localized hydrogen bonds, which locally strengthen the water network. Both aspects have an impact on the polarization of water on the biomolecular surface. Furthermore, the theoretical framework gives a deeper understanding of the assumptions and approximations made in former Lorentzian model [40, 52, 72]. In fact, we can interpret them in terms of the system's behavior on the protein-water interface.

3.3 The Lorentzian model for biomolecules in water

In Section 3.1 we introduced an appropriate representation of the biomolecule, where the polarization effects are captured by a constant dielectric function:

$$\boldsymbol{D}_\Omega(\boldsymbol{r}) = \int_\Omega d\boldsymbol{r}'\, \varepsilon_{\text{local}}(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{E}_\Omega(\boldsymbol{r}'), \qquad \boldsymbol{r}\in\Omega \tag{3.16}$$

$$\text{with} \qquad \varepsilon_{\text{local}}(\boldsymbol{r}-\boldsymbol{r}') = \varepsilon_{\text{macros}}\,\delta(\boldsymbol{r}-\boldsymbol{r}')\,, \qquad \boldsymbol{r},\boldsymbol{r}'\in\Omega\,, \tag{3.17}$$

where Ω is defined as the inside of the molecule, i.e. the region of molecular dielectric response. We agreed on the macroscopic response of the biomolecule, $\varepsilon_{\text{macros}}$ to equal $\varepsilon_\Omega = 2\varepsilon_0$. This finally yields

$$\boldsymbol{D}_\Omega(\boldsymbol{r}) = \varepsilon_\Omega \boldsymbol{E}_\Omega(\boldsymbol{r}), \quad \boldsymbol{r}\in\Omega\,.$$

In Section 3.2 we characterized the dielectric response of water by the Lorentzian model. Here, we assumed the water to capture the whole of $\mathbb{R}^3$. In the biomolecular system, the nonlocal response given in Eqs. (3.14) is confined to $\Sigma\subset\mathbb{R}^3$. With $\varepsilon_{\text{macros}} = \varepsilon_\Sigma = 78.5\,\varepsilon_0$, the macroscopic dielectric response of water, and $\varepsilon_\infty = 1.8\,\varepsilon_0$, its electronic response, we define the nonlocal response in Σ

$$\boldsymbol{D}_\Sigma(\boldsymbol{r}) = \int_\Sigma d\boldsymbol{r}'\, \varepsilon_{\text{nonlocal}}(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{E}_\Sigma(\boldsymbol{r}'), \qquad \boldsymbol{r}\in\Sigma \tag{3.18}$$

$$\text{with} \quad \varepsilon_{\text{nonlocal}}(\boldsymbol{r}-\boldsymbol{r}') = \varepsilon_\infty\delta(\boldsymbol{r}-\boldsymbol{r}') + (\varepsilon_{\text{macros}}-\varepsilon_\infty)\frac{\kappa^2}{4\pi}\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}, \qquad \boldsymbol{r},\boldsymbol{r}'\in\Sigma\,. \tag{3.19}$$

Now, we merge both models in order to define the electrostatic setting of the biomolecular system. To this end, we state a new formulation of the dielectric response in Section 3.3.1, which combines the separate definitions of ε in Ω and Σ to a *global* dielectric operator. Based on this formulation, we discuss the effect of the molecule-water interface.

In Section 3.3.2 we state the system of equations and calculate the field energy of the biomolecular system, so that when we start with first applications in Section 3.3.3, we have a physical quantity to compare with experimental data.

3.3.1 The dielectric function in the biomolecular system

As the material equations are differential equations for $\boldsymbol{E}$ and $\boldsymbol{D}$, which are *globally* defined in $\mathbb{R}^3$, we also need a globally valid dielectric operator $\varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3}$. With the explicitly noted dependence $\boldsymbol{r}$ and $\boldsymbol{r}'$, we already anticipate that the global dielectric operator of the biomolecular system is not homogeneous and isotropic anymore. However, restricted to the two regions Ω and Σ, $\varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3}$ should result in the two separate, homogeneous and isotropic, definitions (3.17) and (3.19), respectively.

We define the projection on a domain $V \in \mathbb{R}^3$ by the characteristic function

$$\chi_V(\boldsymbol{r}) := \begin{cases} 1 & \boldsymbol{r} \in V \\ 0 & \boldsymbol{r} \notin V\,. \end{cases} \tag{3.20}$$

With the help of Eq. (3.20), we define the dielectric permittivity $\boldsymbol{D}$ of the molecular system in $\mathbb{R}^3$

$$\boldsymbol{D}(\boldsymbol{r}) = \int_{\mathbb{R}^3} d\boldsymbol{r}'\, \varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3} \boldsymbol{E}(\boldsymbol{r}'), \quad \boldsymbol{r} \in \mathbb{R}^3\,, \tag{3.21}$$

considering the following response

$$\begin{aligned} \varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3} &= \chi_\Omega(\boldsymbol{r})\chi_\Omega(\boldsymbol{r}')\varepsilon_{\text{local}}(\boldsymbol{r}-\boldsymbol{r}') + \chi_\Sigma(\boldsymbol{r})\chi_\Sigma(\boldsymbol{r}')\varepsilon_{\text{nonlocal}}(\boldsymbol{r}-\boldsymbol{r}') \\ &= \chi_\Omega(\boldsymbol{r})\chi_\Omega(\boldsymbol{r}')\delta(\boldsymbol{r}-\boldsymbol{r}')\varepsilon_\Omega \\ &\quad + \chi_\Sigma(\boldsymbol{r})\chi_\Sigma(\boldsymbol{r}')\left(\varepsilon_\infty\delta(\boldsymbol{r}-\boldsymbol{r}') + (\varepsilon_\Sigma-\varepsilon_\infty)\,\frac{\kappa^2}{4\pi}\,\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\right) \\ &= \varepsilon(\boldsymbol{r}',\boldsymbol{r})_{\mathbb{R}^3}\,. \end{aligned} \tag{3.22}$$

This newly developed global approach is restricted to a homogeneous and isotropic, dielectric response for the cases $(\boldsymbol{r},\boldsymbol{r}') \in \Omega$ *or* $(\boldsymbol{r},\boldsymbol{r}') \in \Sigma$: in Ω, we directly agree with the homogeneous nature, as the dielectric response is a local quantity and therefore the next surrounding of $\boldsymbol{r} \in \Omega$ has no influence on the response at position $\boldsymbol{r}$. In Σ, where the reaction is nonlocal, however, the next surrounding modifies the dielectric response. The restriction to homogeneity and isotropy is actually only valid when water captures the whole space, i.e., $\Sigma = \mathbb{R}^3$ (see Section 3.2.2.1 for further explanations). Thus, in the context of solvated molecules, this is an assumption which implies that the solvent is not perturbed by the biomolecule's cavity, and the dielectric operator reproduces the ideal solvent polarization contribution everywhere in Σ. A direct influence of the dielectric boundary Γ on the equilibrium polarization is then neglected for $\varepsilon_{\text{nonlocal}}$. However, the nonlocal polarization $\boldsymbol{P}_\Sigma$ may dominate the solvent pattern only if the biomolecule "fits well" into the spatial and chemical structure of the water network, i.e., when the hydrogen bonds are not broken or completely rearranged but when they are still smoothly built around the cavity of the molecule (see Section 3.1.4.2, where for the solvation process we introduce entropic energy terms accounting for the symmetry breaking cavity).

This aspect is further explained in Fig. 3.9, where we illustrate three different positions of a water molecule (small black point) near the dielectric boundary Γ. The considered water molecule

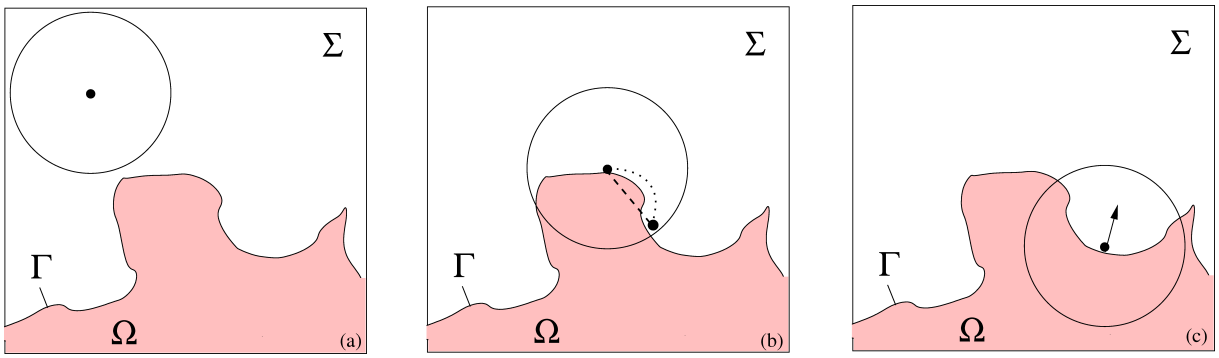


Fig. 3.9: Nonlocal response of the water molecule in its correlation sphere (CS) near the dielectric boundary Γ.

is surrounded by its correlation sphere (CS) which is defined by the correlation length λ. The correlation sphere has a radius of about 2-10Å [33, 124] and therefore "sees" the curvature of the molecule's surface. In Fig. 3.9

(a) the CS of the water molecule under consideration completely lies in Σ. The water molecule is surrounded by other water molecules, where the nonlocal isotropic response causes hydrogen bonds as a function of the local electrostatic field.

(b) the CS penetrates the dielectric boundary in a way that the water network is disturbed. The direct (dashed) connection of the water molecule to another one that lies inside its CS goes through Ω and thus, the water network in this direction is considerably weakened. However, the figure additionally shows another (dotted) connection between the two water molecules. This should indicate that, when (i) the surface is slightly convex and (ii) correlations of the water and the protein's surface exposed amino acids are negligible, the network can be mediated by a rearrangement of other water molecules.

(c) the water molecule lies inside a pocket and its CS is mainly covered by Ω. In surface normal direction (arrow), i.e., the direction of "free" water molecules, the network can be built whereas all the other directions are blocked by the dielectric boundary.

Fig. 3.9 motivates that the nonlocal part of the polarization changes when approaching the dielectric boundary, i.e., when the CS partly lies in Ω: then, the arbitrarily shaped surface weakens the direct water correlations and, in addition, the polarization on Γ might be changed due to local, water-amino acid correlations.

The global, nonlocal dielectric function proposed in Eq. (3.22) in fact "ignores" the dielectric boundary, assuming $\varepsilon(\boldsymbol{r}, \boldsymbol{r}') = \varepsilon(|\boldsymbol{r} - \boldsymbol{r}'|)$ to be the undisturbed nonlocal response when both $\boldsymbol{r}$ and $\boldsymbol{r}'$ lie in Σ – whatever lies *between* $\boldsymbol{r}$ and $\boldsymbol{r}'$. Further, it does not incorporate a possible change to anisotropy for CSs which overlap with Ω. This could be important when incorporating, for instance, local hydrogen bonds with amino acid side chains.

However, the dielectric boundary is *not* ignored in the calculation of the electrostatic fields $\boldsymbol{E}$ and $\boldsymbol{D}$. On the boundary Γ, $\boldsymbol{E}$ has to fulfill transmission conditions which take into account the change in the dielectric response. Moreover, the integration regime in Eq. (3.18),

$$\boldsymbol{D}_\Sigma(\boldsymbol{r}) = \int_\Sigma d\boldsymbol{r}'\, \varepsilon_{\text{nonlocal}}(\boldsymbol{r} - \boldsymbol{r}')\boldsymbol{E}_\Sigma(\boldsymbol{r}')\,, \qquad \boldsymbol{r} \in \Sigma\,,$$

explicitly delimits the nonlocal response to Σ. This implies an indirect incorporation of the dielectric boundary and its shape when calculating the dielectric permittivity $\boldsymbol{D}$.

As a first approach, we assume that any direct variations in Eq. (3.22) due to the dielectric boundary are negligible compared to the indirect influence the dielectric boundary has on $\boldsymbol{D}$. In Section 4.3.2, we once again refer to the question how to incorporate boundary effects and find a way to easily incorporate a predefined behavior of the nonlocal response on Γ.

3.3.2 The Maxwell equations of the biomolecular system

Our aim is to analyze how the nonlocal Lorentzian solvent impacts the electrostatics and finally the characteristics of solvated biomolecules. We therefore write the globally valid material equations for a molecule in water analogous to Theorem 2.2.1,

$$\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{D} &= -\rho, && \text{in } \mathbb{R}^3 \\ \boldsymbol{\nabla}\times\boldsymbol{E} &= \boldsymbol{0}, && \text{in } \mathbb{R}^3\,, \end{aligned}$$

and, based on the following definitions, cast these global equations into a transmission problem using Eqs. (3.21) and (3.22).

Definition 3.1

$$\begin{aligned} \boldsymbol{F}_\Sigma &= \frac{\kappa^2}{4\pi}\int\limits_\Sigma d\boldsymbol{r}'\,\frac{e^{-\kappa\,|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\boldsymbol{E}_\Sigma(\boldsymbol{r}'), && \text{in } \Sigma \qquad (3.23)\\ \boldsymbol{E}_\Omega &= -\boldsymbol{\nabla}\phi_\Omega, && \text{in } \Omega \\ \boldsymbol{E}_\Sigma &= -\boldsymbol{\nabla}\phi_\Sigma, && \text{in } \Sigma \\ \boldsymbol{D}_\Omega &= -\varepsilon_\Omega\boldsymbol{\nabla}\phi_\Omega, && \text{in } \Omega \\ \boldsymbol{D}_\Sigma &= -\left(\varepsilon_\infty\boldsymbol{\nabla}\phi_\Sigma - (\varepsilon_\Sigma-\varepsilon_\infty)\,\boldsymbol{F}_\Sigma\right), && \text{in } \Sigma \end{aligned}$$

Theorem 3.3.1 *The Lorentzian model for a molecule in water in integro-differential form: With Definition 3.1, the integro-differential system of nonlocal electrostatics in the Lorentzian model reads*

$$\begin{aligned} \varepsilon_\Omega\triangle\phi_\Omega &= -\rho, && \text{in } \Omega \\ \varepsilon_\infty\triangle\phi_\Sigma - (\varepsilon_\Sigma-\varepsilon_\infty)\,\boldsymbol{\nabla}\cdot\boldsymbol{F}_\Sigma &= 0, && \text{in } \Sigma \\ \varepsilon_\Omega\partial_n\phi_\Omega - \left(\varepsilon_\infty\partial_n\phi_\Sigma - (\varepsilon_\Sigma-\varepsilon_\infty)\,(\boldsymbol{n}\cdot\boldsymbol{F}_\Sigma)\right) &= \sigma, && \text{on } \Gamma \\ \phi_\Omega - \phi_\Sigma &= 0, && \text{on } \Gamma \end{aligned}$$

Remark 3.2 The system in Theorem 3.3.1 is not restricted to a charge distribution in Ω but allows for a surface charge distribution σ on Γ. This is to capture the analytically treatable examples in Section 3.3.3. For the study of biomolecules, we always set σ to 0, see Section 3.1.2 for the surface definition.

3.3.2.1 The correlation field

In Theorem 3.3.1 we have introduced the vector field

$$\boldsymbol{F}_\Sigma(\boldsymbol{r}) = \frac{\kappa^2}{4\pi}\int\limits_\Sigma d\boldsymbol{r}'\,\frac{e^{-\kappa\,|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\boldsymbol{E}_\Sigma(\boldsymbol{r}'), \qquad \boldsymbol{r}\in\Sigma\,.$$

It serves to decompose the dielectric permittivity into a local and a nonlocal part in Σ:

$$\begin{aligned} \boldsymbol{D}_\Sigma(\boldsymbol{r}) &= \varepsilon_0 \boldsymbol{E}_\Sigma(\boldsymbol{r}) + \boldsymbol{P}_\Sigma(\boldsymbol{r}) \\ &= \underbrace{\varepsilon_\infty \boldsymbol{E}_\Sigma(\boldsymbol{r})}_{\text{local dependence}} + \underbrace{(\varepsilon_\Sigma - \varepsilon_\infty)\boldsymbol{F}_\Sigma(\boldsymbol{r})}_{\text{nonlocal dependence}}, \quad \boldsymbol{r} \in \Sigma \end{aligned}$$

In general, the field $\boldsymbol{F}_\Sigma$ comprises the nonlocal part of the polarization $\boldsymbol{P}_\Sigma$, which originates from the orientational degrees of freedom of the water molecules and its correlations. The difference of the field $\boldsymbol{F}_\Sigma$ and the polarization field $\boldsymbol{P}_\Sigma$ originates from the internal, electronic degrees of freedom of every individual water molecule characterized by the constant response ε_∞. If we assume the effect of this electronic polarization to be small, i.e., $\varepsilon_\infty \approx \varepsilon_0$, the field $\boldsymbol{F}_\Sigma$ equals the polarization field $\boldsymbol{P}_\Sigma$. In the following, we denote $\boldsymbol{F}_\Sigma$ as the *correlation field*, having in mind that it takes into account only the *nonlocal part* of $\boldsymbol{P}_\Sigma$.

From the definition of the polarization field $\boldsymbol{F}_\Sigma$ we can already extract two limiting values for low λ (high κ) and high λ (low κ),

$$\lim_{\kappa \to 0} \boldsymbol{F}_\Sigma = \boldsymbol{0} \qquad \Rightarrow \qquad \boldsymbol{D} = \varepsilon_\infty \boldsymbol{E}, \quad \text{in} \Sigma \tag{3.24}$$

$$\lim_{\kappa \to \infty} \boldsymbol{F}_\Sigma = \boldsymbol{E}_\Sigma \qquad \Rightarrow \qquad \boldsymbol{D} = \varepsilon_\Sigma \boldsymbol{E}, \quad \text{in} \Sigma\ . \tag{3.25}$$

From Eq. (3.24) we read that the correlation field vanishes when the strength of the hydrogen bonds between the water molecules becomes maximal. This freezes the dielectric response to a pure electronic response, because the water molecules cannot align their permanent dipoles anymore. The nonlocal model is then described by a local model with electronic response ε_∞ in Σ.

Eq. (3.25) tells us that when the correlation length vanishes, the water molecules are free to align their permanent dipole with the dielectric field $\boldsymbol{D}$ and we regain the macroscopic dielectric response ε_{water}. This means that we end up in the expected local model with dielectric response ε_Σ in Σ.

Assuming an intermediate value of the correlation length, the correlation field describes a response lying between these two limiting values, which are both represented by local models.

In Section 3.3.1, we discussed the problem of a possible change of the nonlocal response near the dielectric boundary Γ. We have motivated that for a first model development of nonlocal electrostatics, it is reasonable to neglect a direct, local change of $\varepsilon(\boldsymbol{r}, \boldsymbol{r}')_{\mathbb{R}^3}$ at Γ and to assume an isotropic, homogeneous dielectric operator as given by Eq. (3.22). Since this problem only affects the part of the polarization that accounts for the nonlocal correlation, the decision to keep a homogeneous, isotropic dielectric function is expressed by the features of $\boldsymbol{F}_\Sigma$.

A closer look at Eq. (3.23) tells us that the correlation field is defined in the whole of $\mathbb{R}^3$:

$$\boldsymbol{F}(\boldsymbol{r}) = \frac{\kappa^2}{4\pi} \int_\Sigma d\boldsymbol{r}' \frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} \boldsymbol{E}_\Sigma(\boldsymbol{r}'), \qquad \boxed{\boldsymbol{r} \in \mathbb{R}^3}. \tag{3.26}$$

This, in turn, determines the boundary values of the correlation field and therefore the behavior of the nonlocal correlations on Γ. A homogeneous, isotropic dielectric operator in Σ implies that the *extension* of $\boldsymbol{F}_\Sigma$ in $\mathbb{R}^3$ is continuous with continuous normal derivatives. Indeed, one can prove that the correlation field $\boldsymbol{F}$ in Eq. (3.26) is continuous and every component of $\boldsymbol{F}$ has a continuous normal derivative (it is a Newton potential of the Yukawa operator [121]).

The correlation field $\boldsymbol{F}$ has been first defined in [39, 40]. However, its physical interpretation as given above and a discussion of reasonable modifications has not yet been the focus in former studies. In Section 4.3, this is once again addressed.

3.3.2.2 The field energy of the biomolecular system

In Section 2.3 we found that in order to obtain an explicit expression for the field energy W_{field} in dielectric media, we have to specify the dielectric operator. We proved that the field energy of systems with piecewise linear, local response is given by

$$W_{field} = \frac{1}{2}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \phi(\boldsymbol{r})\, \rho(\boldsymbol{r}) = \frac{1}{2}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \boldsymbol{D}(\boldsymbol{r}) \cdot \boldsymbol{E}(\boldsymbol{r})\,.$$

In this section, we calculate the field energy W_{field} for the Lorentzian model of a solvated molecule. With the global, nonlocal ansatz (Eqs. (3.22) and (3.21))

$$\boldsymbol{D}(\boldsymbol{r}) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, \varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3} \boldsymbol{E}(\boldsymbol{r}')\,,$$

the infinitesimal work δW_{field} caused by adding an increment of charge density $\delta\rho$ to each volume element $d\boldsymbol{r}$ is given by

$$\delta W_{field} = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \phi(\boldsymbol{r})\delta\rho(\boldsymbol{r}) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \boldsymbol{E}(\boldsymbol{r}) \cdot \delta\boldsymbol{D}(\boldsymbol{r}) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \boldsymbol{E}(\boldsymbol{r}) \cdot \left(\int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, \varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3}\, \delta\boldsymbol{E}(\boldsymbol{r}') \right)$$

The integrals commute, because

$$\varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3} = \varepsilon(\boldsymbol{r}',\boldsymbol{r})_{\mathbb{R}^3}\,,$$

and thus we can shift the variation upon $\boldsymbol{E}$:

$$\begin{aligned}
\delta W_{field} &= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \delta\boldsymbol{D}(\boldsymbol{r}) \cdot \boldsymbol{E}(\boldsymbol{r}) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, \boldsymbol{D}(\boldsymbol{r}') \cdot \delta\boldsymbol{E}(\boldsymbol{r}') \\
&= \frac{1}{2}\delta \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \boldsymbol{D}(\boldsymbol{r}) \cdot \boldsymbol{E}(\boldsymbol{r}) = \frac{1}{2}\delta\langle \boldsymbol{D}(\boldsymbol{r}), \boldsymbol{E}(\boldsymbol{r})\rangle \\
W_{field} &= \frac{1}{2} \int\limits_0^{\langle \boldsymbol{D},\boldsymbol{E}\rangle} \delta\langle \boldsymbol{D}(\boldsymbol{r}), \boldsymbol{E}(\boldsymbol{r})\rangle = \frac{1}{2}\langle \boldsymbol{D}, \boldsymbol{E}\rangle \\
&= \frac{1}{2}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \boldsymbol{D}(\boldsymbol{r}) \cdot \boldsymbol{E}(\boldsymbol{r}) = \frac{1}{2}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \phi(\boldsymbol{r})\, \rho(\boldsymbol{r})
\end{aligned}$$

This means that even for a piecewise (non)local response, the common formula for the electrostatic field energy is valid. For its derivation, the requirements on linearity of the dielectric response and the commutability of $\boldsymbol{r}$ and $\boldsymbol{r}'$ in the dielectric operator are necessary and sufficient. We end up with Theorem 3.3.2.

Theorem 3.3.2 *The electrostatic field energy for linear, piecewise, (non)local dielectric media in $\mathbb{R}^3$: If the dielectric response of the medium in $\mathbb{R}^3$ is given by the following equations,*

$$\boldsymbol{D}(\boldsymbol{r}) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, \varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3} \boldsymbol{E}(\boldsymbol{r}')$$

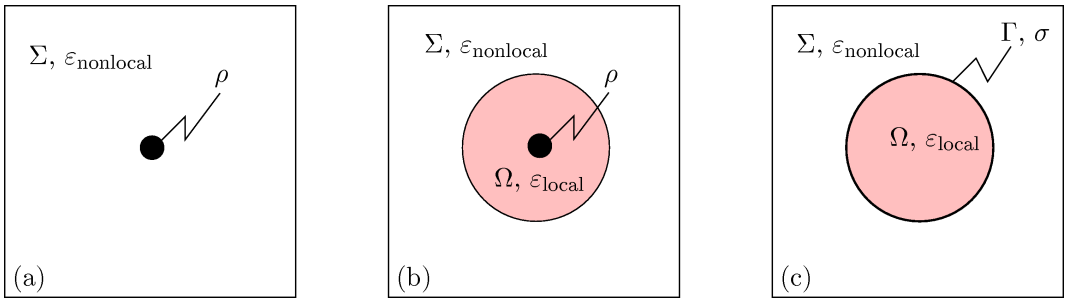


Fig. 3.10: Different setups with spherical symmetry, where the piecewise (non)local dielectric function is defined in (Eq. (3.19)) Eq. (3.17).

$$\varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3} = \chi_\Omega(\boldsymbol{r})\chi_\Omega(\boldsymbol{r}')\delta(\boldsymbol{r}-\boldsymbol{r}')\varepsilon_\Omega + \chi_\Sigma(\boldsymbol{r})\chi_\Sigma(\boldsymbol{r}')\left(\varepsilon_\infty\delta(\boldsymbol{r}-\boldsymbol{r}') + (\varepsilon_\Sigma-\varepsilon_\infty)\frac{\kappa^2}{4\pi}\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\right),$$

the electrostatic field energy, W_{field} of the system with charge distribution ρ is given by

$$W_{field} \;=\; \frac{1}{2}\int_{\mathbb{R}^3} d\boldsymbol{r}\, \boldsymbol{D}(\boldsymbol{r})\cdot\boldsymbol{E}(\boldsymbol{r}) = \frac{1}{2}\int_{\mathbb{R}^3} d\boldsymbol{r}\,\phi(\boldsymbol{r})\,\rho(\boldsymbol{r})\,.$$

Important remark 3.2 In Section 4.3 the definition of the correlation field, Eq. (3.23),

$$\boldsymbol{F}_\Sigma(\boldsymbol{r}) = \frac{\kappa^2}{4\pi}\int_\Sigma d\boldsymbol{r}'\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\boldsymbol{E}_\Sigma(\boldsymbol{r}'), \qquad \boldsymbol{r}\in\Sigma\,,$$

is extended to two further boundary integrals. As we will address in detail in Remark 4.4, an incorporation of these boundary integrals into the calculation of the field energy as described above is difficult. For the present work, we always calculate the energy formula as given in Theorem 3.3.2, having in mind that this is an approximation for general Lorentzian models of nonlocal response.

3.3.3 Application to spherically symmetric systems

In Fig. 3.10 we illustrate three different settings, which we want to solve in a nonlocal framework. Since all the setups possess spherical symmetry, the solutions will be spherically symmetric and this makes an analytical treatment of the integro-differential system (Theorem 3.3.1) possible.

Before we go into detail, we recapitulate two properties of radially symmetric fields. First, a spherically symmetric, differentiable vector field $\boldsymbol{\mathcal{F}}$ can be identified by $\boldsymbol{\mathcal{F}}(\boldsymbol{r}) = \boldsymbol{e}_r f(r)$, where $\boldsymbol{e}_r$ defines the unity vector in direction $\boldsymbol{r}$ with $r = |\boldsymbol{r}|$ and f is a differentiable scalar function. Second, for a radially symmetric field one finds a representation [5],

$$\boldsymbol{\mathcal{F}} = \boldsymbol{e}_r f(r) = -\boldsymbol{\nabla}\mathcal{F}(r)\;, \quad \text{where } f(r) = -\partial_r\mathcal{F}(r)\,.$$

Thus, in addition to the gradient representation of the electrostatic field, $\boldsymbol{E} = -\boldsymbol{\nabla}\phi$, we find a potential representation of the dielectric permittivity $\boldsymbol{D}$ as well as of the correlation field $\boldsymbol{F}$ in the spherically symmetric setting. We introduce the *dielectric potential* ψ and the potential F and define

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi\,, \qquad \boldsymbol{D} = -\boldsymbol{\nabla}\psi\,, \qquad \boldsymbol{F} = -\boldsymbol{\nabla}F\,.$$

[5] The rotation of a radially symmetric fields vanishes and this directly implies a gradient representation.

3.3.3.1 The point charge

The simplest application of nonlocal electrostatics is shown in Fig. 3.10(a), where a point charge is immersed in the nonlocal solvent. In this case, the dielectric function, $\varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3}$ defined in Eq. (3.22), is reduced to the globally defined, nonlocal dielectric operator

$$\varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3} = \varepsilon_{\text{nonlocal}}(\boldsymbol{r}-\boldsymbol{r}') = \varepsilon_\infty \delta(\boldsymbol{r}-\boldsymbol{r}') + (\varepsilon_\Sigma - \varepsilon_\infty)\,\frac{\kappa^2}{4\pi}\frac{e^{-\kappa\,|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}, \quad \boldsymbol{r},\boldsymbol{r}' \in \mathbb{R}^3\,.$$

The differential system of electrostatics for the point charge in water then reads

$$\begin{aligned} \triangle\psi(\boldsymbol{r}) &= -q\,\delta(\boldsymbol{r}), && \boldsymbol{r}\in\mathbb{R}^3 \\ \boldsymbol{\nabla}\psi(\boldsymbol{r}) &= (\varepsilon_{nonlocal} * \boldsymbol{\nabla}\phi)(\boldsymbol{r}), && \boldsymbol{r}\in\mathbb{R}^3\,. \end{aligned}$$

Plugging the second into the first equation gives an unique equation for ϕ, which is, in a second step, transformed into Fourier space:

$$\begin{aligned} & -\boldsymbol{\nabla}\left((\varepsilon_{nonlocal} * \boldsymbol{\nabla}\phi)(\boldsymbol{r})\right) = q\,\delta(\boldsymbol{r})\,, && \boldsymbol{r}\in\mathbb{R}^3 \\ \Rightarrow \qquad & \boldsymbol{k}\left(\hat{\varepsilon}_{nonlocal}(k)\,\boldsymbol{k}\hat{\phi}(k)\right) = \frac{q}{\sqrt{2\pi}^3}\,, && k=|\boldsymbol{k}|, \boldsymbol{k}\in\mathbb{R}^3 \end{aligned}$$

with Eq. (3.10) $$\hat{\phi}(k) = \frac{q}{\sqrt{2\pi}^3}\frac{1}{\varepsilon_\Sigma k^2}\frac{1+k^2/\kappa^2}{1+\lambda^2 k^2}, \qquad k=|\boldsymbol{k}|, \boldsymbol{k}\in\mathbb{R}^3$$

Finally, we transform back $\hat{\phi}$ into real space:

$$\phi(r) = \frac{q}{4\pi r}\left(\frac{1}{\varepsilon_\Sigma} + (\frac{1}{\varepsilon_\Sigma} - \frac{1}{\varepsilon_\infty})e^{-r/\lambda}\right), \quad r=|\boldsymbol{r}|, \boldsymbol{r}\in\mathbb{R}^3 \tag{3.27}$$

This nonlocal influence can be directly seen in Eq. (3.27): besides the well known local term $\left(\frac{q}{4\pi\,\varepsilon_\Sigma\,r}\right)$, there is an additional, exponentially decreasing contribution,

$$\frac{q}{4\pi r}\left((\frac{1}{\varepsilon_\Sigma} - \frac{1}{\varepsilon_\infty})e^{-r/\lambda}\right) > 0\,,$$

which causes the potential to reach farther in space (see Fig. 3.11 on p. 49). This is reasonable, since the water network hinders the individual water molecules to align with the field and therefore weakens the macroscopic response that is given by ε_Σ. The nonlocal effect has its maximum next to the point charge and decreases on the scale of the correlation length λ.

As the nonlocal part of ϕ decreases faster than the local one, the electrostatic potential results in the known expression for the local potential far from the point charge. A detailed comparison between the local and the nonlocal behavior of the potential ϕ is given in Section 3.3.3.4.

In Theorem 2.1.2 we already used the linearity of the differential equations to generalize the solution of the Laplace equation with source term $q\,\delta(\boldsymbol{r}-\boldsymbol{r}')$ to an arbitrary charge distribution. The same can be done here and we end up with Theorem 3.3.3.

Theorem 3.3.3 *Point charges immersed in water: An arbitrary distribution of N point charges*

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N} q_i \delta(\boldsymbol{r}-\boldsymbol{r}_i)$$

is immersed in a solvent of Lorentzian response. The electrostatic potential ϕ*,*

$$\phi(\boldsymbol{r}) = \frac{1}{4\pi}\sum_{i=1}^{N}\frac{q_i}{|\boldsymbol{r}-\boldsymbol{r}_i|}\left(\frac{1}{\varepsilon_\Sigma}+(\frac{1}{\varepsilon_\Sigma}-\frac{1}{\varepsilon_\infty})e^{-\frac{1}{\lambda}|\boldsymbol{r}-\boldsymbol{r}_i|}\right),$$

solves the global, electrostatic material equation of Lorentzian response,

$$\boldsymbol{\nabla}(\varepsilon_{nonlocal} * \boldsymbol{\nabla}\phi)(\boldsymbol{r}) = -\rho(\boldsymbol{r}), \quad \boldsymbol{r}\in\mathbb{R}^3.$$

Remark 3.3 The potential ϕ can be expressed by the fundamental solution $\boldsymbol{G}_{\mathcal{D}}$,

$$\begin{array}{rccl} \boldsymbol{G}_{\mathcal{D}}: & \mathbb{R}^3\backslash\{0\} & \mapsto & \mathbb{R}^2 \\ & \boldsymbol{r} & \mapsto & \begin{pmatrix} \frac{1}{4\pi|\boldsymbol{r}|} \\ \frac{1}{4\pi|\boldsymbol{r}|}\left(\frac{1}{\varepsilon_\Sigma}+(\frac{1}{\varepsilon_\Sigma}-\frac{1}{\varepsilon_\infty})e^{-\frac{1}{\lambda}|\boldsymbol{r}|}\right) \end{pmatrix} \end{array}$$

of the differential system with operator $\mathcal{D}$ (see Theorem 2.1.2 and Section 4 for further explanations):

$$\mathcal{D} = \begin{pmatrix} \triangle & 0 \\ \frac{1}{\lambda^2}\frac{\varepsilon_\infty}{\varepsilon_\Sigma} & \varepsilon_\infty(\triangle - \frac{1}{\lambda^2}) \end{pmatrix}.$$

3.3.3.2 The Born sphere

The Born model has been introduced in Section 3.1.1. It is the simplest model for monoatomic ions. Studying the Born model is an important benchmark for nonlocal electrostatics, because in our analysis, molecules are treated as an accumulation of Born spheres. Now, we want to solve the nonlocal electrostatics of the Born sphere.

Fig. 3.10(b) sketches the situation: Ω equals a sphere of arbitrary radius a, the surface charge is set to zero and the charge distribution is the one of a point charge $\rho = q\delta(\boldsymbol{r})$. In order to represent a monoatomic ion, Ω is filled with vacuum and therefore the response inside the sphere is $\varepsilon_\Omega = \varepsilon_0$. Since our derivation is valid for arbitrary constant dielectric response, we do not specify ε_Ω further.

As the global definition of the dielectric function,

$$\varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3} = \chi_\Omega(\boldsymbol{r})\chi_\Omega(\boldsymbol{r}')\,\varepsilon_\Omega\delta(\boldsymbol{r}-\boldsymbol{r}') + \chi_\Sigma(\boldsymbol{r})\chi_\Sigma(\boldsymbol{r})\left(\varepsilon_\infty\delta(\boldsymbol{r}-\boldsymbol{r}') + (\varepsilon_\Sigma-\varepsilon_\infty)\frac{\kappa^2}{4\pi}\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\right),$$

cannot be reduced to a function of $\varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3} = \varepsilon(\boldsymbol{r}-\boldsymbol{r}')_{\mathbb{R}^3}$ as it was possible for the point charge, we cannot use the convolution theorem to solve the nonlocal Maxwell equations [6]. This is why we solve the system for the Born sphere in real space:

$$\left.\begin{array}{lll} \boldsymbol{F}_\Sigma = -\boldsymbol{\nabla}F_\Sigma(\boldsymbol{r}) = -\frac{\kappa^2}{4\pi}\int\limits_\Sigma d\boldsymbol{r}'\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\boldsymbol{\nabla}_{r'}\phi_\Sigma(\boldsymbol{r}'), & & r>a \\ \varepsilon_\Omega\triangle\phi_\Omega & = -q\delta(\boldsymbol{r}), & r<a \\ \triangle\underbrace{(\varepsilon_\infty\phi_\Sigma + (\varepsilon_\Sigma-\varepsilon_\infty)F_\Sigma)}_{\psi_\Sigma} & = 0, & r>a \\ \phi_\Omega - \phi_\Sigma & = 0, & r=a \\ \varepsilon_\Omega\partial_n\phi_\Omega - \partial_n\underbrace{(\varepsilon_\infty\phi_\Sigma + (\varepsilon_\Sigma-\varepsilon_\infty)F_\Sigma)}_{\psi_\Sigma} & = 0, & r=a \end{array}\right\} \quad (3.28)$$

[6] Such an approach is taken in [52] and therefore the solutions given therein are slightly wrong.

In order to solve the system (3.28) we need an ansatz for the potentials $(\phi_\Sigma, \phi_\Omega, \psi_\Sigma)$. In Section 2.1.1 we discussed the general solution of the Laplace operator for spherically symmetric systems. In addition, we suggest that the principal behavior of ϕ_Σ corresponds to the one of the point charge, Eq. (3.27) [7]. From this, we make the following approach:

$$\begin{aligned} \phi_\Sigma &= \frac{1}{4\pi r}\left(\frac{1}{\varepsilon_\Sigma} + B\, e^{-s(r-a)}\right) \\ \phi_\Omega &= \frac{1}{4\pi}\left(\frac{1}{\varepsilon_\Omega r} + A\right) \\ \psi_\Sigma &= \frac{1}{4\pi\, r} \end{aligned} \tag{3.29}$$

Because of the spherical symmetry, we switch to spherical coordinates (r, φ, ϑ). Without loss of generality we assume the considered point $\boldsymbol{r}$ to lie on the z-axis, i.e., $\boldsymbol{r} = (0, 0, r)$, $r > a$.

$$\begin{aligned} \boldsymbol{F}_\Sigma(\boldsymbol{r}) &= -\frac{\kappa^2}{4\pi}\int_a^\infty dr' r'^2 \partial_{r'}\phi_\Sigma(r') \int_0^\pi d\vartheta \sin\vartheta \frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} \int_0^{2\pi} d\varphi\, (\cos\varphi\, \sin\vartheta, \sin\varphi\, \cos\vartheta, \cos\vartheta)^T \\ &= -\frac{\kappa^2}{2}\int_a^\infty dr' r'^2 \partial_{r'}\phi_\Sigma(r') \int_1^{-1} d\cos\vartheta \frac{e^{-\kappa\sqrt{r^2+r'^2-2\,r\,r'\cos\vartheta}}}{\sqrt{r^2+r'^2-2\,r\,r'\cos\vartheta}} (0, 0, \cos\vartheta)^T \\ &= -\frac{1}{2\,\kappa\, r^2}\int_a^\infty dr' \partial_{r'}\phi_\Sigma(r') \Big((\kappa^2 r\, r' + \kappa(r+r') + 1) e^{-\kappa(r+r')} + \\ &\qquad (\kappa^2 r\, r' - \kappa|r-r'| - 1) e^{-\kappa(|r-r'|)}\Big)(0, 0, 1)^T,\ r > a \end{aligned} \tag{3.30}$$

For a further evaluation, we replace in Eq. (3.30) ϕ_Σ by the ansatz given in Eq. (3.29). The integration can then be carried out and yields

$$\begin{aligned} F_\Sigma(r) &= \frac{1}{4\pi\varepsilon_\Sigma\, r} + B\frac{e^{-s(r-a)}}{r}\frac{\kappa^2}{\kappa^2 - s^2} \\ &+ \frac{e^{-\kappa r}}{\kappa a r}\left(\sinh(\kappa a) + B\frac{\cosh(\kappa a)\kappa s^2 a + \sinh(\kappa a)(\kappa^2 s a + \kappa^2 - s^2)}{\kappa^2 - s^2}\right),\ r > a\,, \end{aligned} \tag{3.31}$$

where, for the sake of clarity, we wrote the scalar potential F_Σ. $\boldsymbol{F}_\Sigma$ is regained by applying the gradient operator, i.e., $\boldsymbol{F}_\Sigma = -\boldsymbol{\nabla} F_\Sigma$.

In the last step, the transmission conditions of system (3.28) have to be fulfilled and the relation

$$\frac{1}{4\pi\, r} = \psi_\Sigma := \varepsilon_\infty \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty) F_\Sigma, \qquad \forall \boldsymbol{r} \in \Sigma\,, \tag{3.32}$$

has to be valid. This determines the three constants (A, B, s):

$$\begin{aligned} s &= \frac{1}{\lambda} = \sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}\kappa \\ A &= \left(\frac{1}{\varepsilon_\Sigma} - \frac{1}{\varepsilon_\Omega} + B\right)\frac{1}{a} \\ B &= -\frac{\sinh(\kappa a)(\kappa^2 - s^2)}{\cosh(\kappa a)\kappa s^2 a + \sinh(\kappa a)(\kappa^2 s a + \kappa^2 - s^2)} = \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma}\right)\frac{1}{1 + s\,a + \frac{\varepsilon_\Sigma}{\varepsilon_\infty}\left(\frac{\cosh(\kappa\, a)\,\kappa\, a}{\sinh(\kappa\, a)} - 1\right)} \end{aligned}$$

[7] *A posteriori* this is justified, as we find a solution. In Section 4.3.1.1 we state a purely differential system equivalent to (3.28) with the same, *unique* solution.

For the limit $a \to 0$ the Born potential ϕ_Σ turns into the potential of a point charge as it was derived in the last section. The limit $\kappa \to \infty$ is realized by $B \to 0$ and corresponds to the well known result for a local macroscopic response characterized by ε_Σ.

3.3.3.3 The spherical shell

An illustration of the spherical shell is given in Fig. 3.10(c). Ω equals a sphere of radius a, the surface has a constant charge distribution σ, and the charge density inside the sphere equals zero. For the spherical shell the system given in Theorem 3.3.1 simplifies to

$$\left.\begin{aligned}
&\boldsymbol{F}_\Sigma = -\boldsymbol{\nabla} F_\Sigma = -\tfrac{\kappa^2}{4\pi} \int\limits_\Sigma d\boldsymbol{r}' \tfrac{e^{-\kappa\,|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} \boldsymbol{\nabla}_{r'} \phi_\Sigma(\boldsymbol{r}'), && && r > a \\
&\varepsilon_\Omega \triangle \phi_\Omega && = && 0, \quad r < a \\
&\triangle \underbrace{\left(\varepsilon_\infty \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty) F_\Sigma\right)}_{\psi_\Sigma} && = && 0, \quad r > a \\
&\phi_\Omega - \phi_\Sigma && = && 0, \quad r = a \\
&\varepsilon_\Omega \partial_n \phi_\Omega - \partial_n \underbrace{\left(\varepsilon_\infty \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty) F_\Sigma\right)}_{\psi_\Sigma} && = && \sigma, \quad r = a
\end{aligned}\right\} \tag{3.33}$$

As the potential ϕ_Ω lies in the kernel of the Laplace operator, we make the following ansatz:

$$\begin{aligned}
\phi_\Sigma &= \tfrac{1}{4\pi\, r}\left(\tfrac{1}{\varepsilon_\Sigma} + B\, e^{-s(r-a)}\right) \\
\phi_\Omega &= \tfrac{1}{4\pi} A \\
\psi_\Sigma &= \tfrac{1}{4\pi\, r}
\end{aligned}$$

It remains to determine the constants (A, B, s) with Eq. (3.32) and the transmission conditions of the system (3.33):

$$\begin{aligned}
s &= \tfrac{1}{\lambda} = \sqrt{\tfrac{\varepsilon_\Sigma}{\varepsilon_\infty}}\kappa \\
A &= \left(\tfrac{1}{\varepsilon_\Sigma} - \tfrac{1}{\varepsilon_\Omega} + B\right)\tfrac{1}{a} \\
B &= -\frac{\sinh(\kappa a)(\kappa^2 - s^2)}{\cosh(\kappa a)\kappa s^2 a + \sinh(\kappa a)(\kappa^2 s a + \kappa^2 - s^2)} \\
&= \left(\tfrac{1}{\varepsilon_\infty} - \tfrac{1}{\varepsilon_\Sigma}\right) \frac{1}{1 + s\,a + \frac{\varepsilon_\Sigma}{\varepsilon_\infty}\left(\frac{\cosh(\kappa\, a)\,\kappa\, a}{\sinh(\kappa\, a)} - 1\right)}
\end{aligned}$$

3.3.3.4 Discussion and comparison

When comparing the Born with the spherical shell model, we see that the potentials *outside* the sphere are equivalent. This is valid separately for both, the local and the nonlocal dielectric response. It is a consequence of the physical Gauss law, Eq. (2.3), which tells us that the dielectric potential outside a sphere is equivalent for all spherically symmetric charge distributions enclosed in the sphere with the same total charge.

Fig. 3.11 shows the electrostatic potential ϕ of the Born model (lined), of the shell model (dashed), and of the point charge (dotted). The difference to the point charge can be seen inside the sphere, but it is not resolved anymore outside the sphere.

Compared to the local theory, the nonlocal models have a higher electrostatic potential everywhere. This is in accordance with the fact that the water network hinders the individual water molecules to freely rotate and, in this way, the network shields the electrostatic field. All models end up in the same local limit for $r \to \infty$ (not shown). This corresponds to the constraint on the nonlocal theory to turn into the local macroscopic behavior when the field variations get small.

The discontinuity of the normal derivative of the potential can also be seen in Fig. 3.11. For the

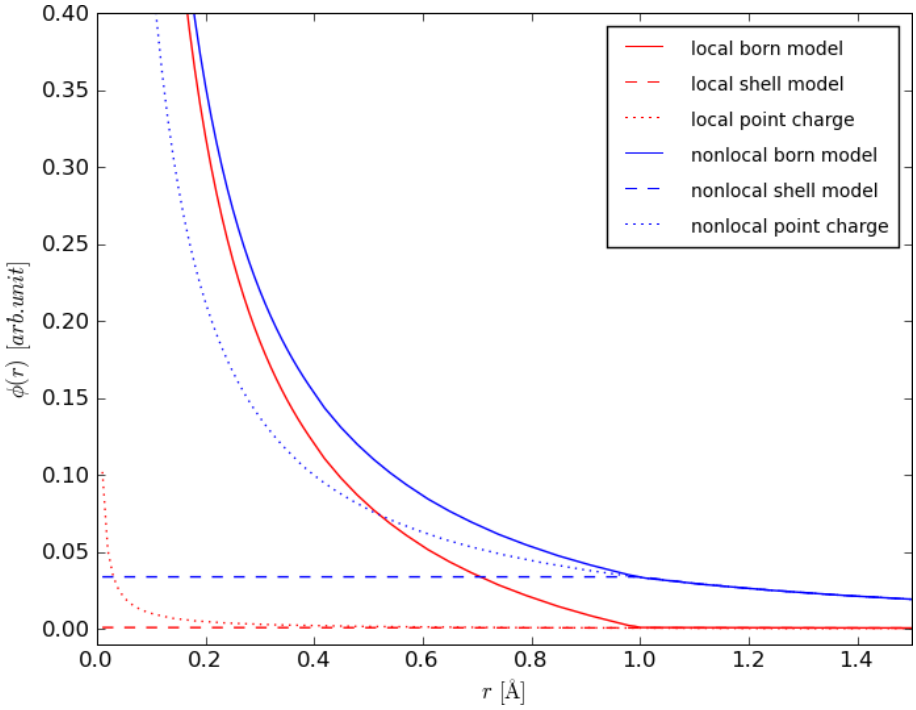


Fig. 3.11: The shape of electrostatic potential for different spherically symmetric cases, $\kappa = 1/23$ Å^{-1}, $\varepsilon_\Omega = 1\,\varepsilon_0$, $\varepsilon_\Sigma = 78.0\,\varepsilon_0$, $\varepsilon_\infty = 1.8\,\varepsilon_0$, $a = 1$Å, thus the distance $r = 1$Å corresponds to Γ.

Born sphere, it is exclusively caused by a change in the dielectric function, whereas for the shell model there is a further contribution due to the non-vanishing surface charge σ. The jump at Γ for the nonlocal Born model can hardly be seen. Only in comparison with the continuous function of the nonlocal point charge (dotted), we recognize that there is a change in the slope at Γ. In contrast, the local model (lined functions) causes a change in the slope of $\sim\varepsilon_\Sigma/\varepsilon_\Omega$, as can be easily seen in the figure. This is an interesting result, because the "smooth" transition of the potential predicted by nonlocal model implies that the macroscopic response characterized by ε_Σ is *indirectly* weakened on the surface. This was already supposed in Section 3.3.1 when we discussed a possible change of the dielectric operator because of the dielectric boundary.

In the last decade, much attention has been given to the experimental determination and theoretical calculation of the free solvation energies or free enthalpies to describe the transfer of individual ions between two solvents. The interest in this quantity is due to its crucial importance in bioenergetics and charge transfer processes in electrolyte solutions.

With the background derived in Section 3.1.4, we can apply the nonlocal theory to calculate the electrostatic contribution to the solvation free energies. In Section 3.1.1, we introduced the Born and the shell model as useful representations of monoatomic ions. Using these models, the complexity of their quantum mechanical nature is not captured and therefore, in order to approximate experimentally measurable quantities, one usually takes the atom radius as a fit parameter. However, to assess the performance of a *new* theory we want to avoid a fitting process. In [52] A. Hildebrandt explains in detail how physically motivated radii can be extracted from radial distribution functions based on the ideas of Åqvist. In the following application of the nonlocal model to calculate solvation energies, we represent monoatomic metal ions by Born spheres with radii taken from [1].

In [33, 124] and references given therein, the correlation length λ has been assumed to range

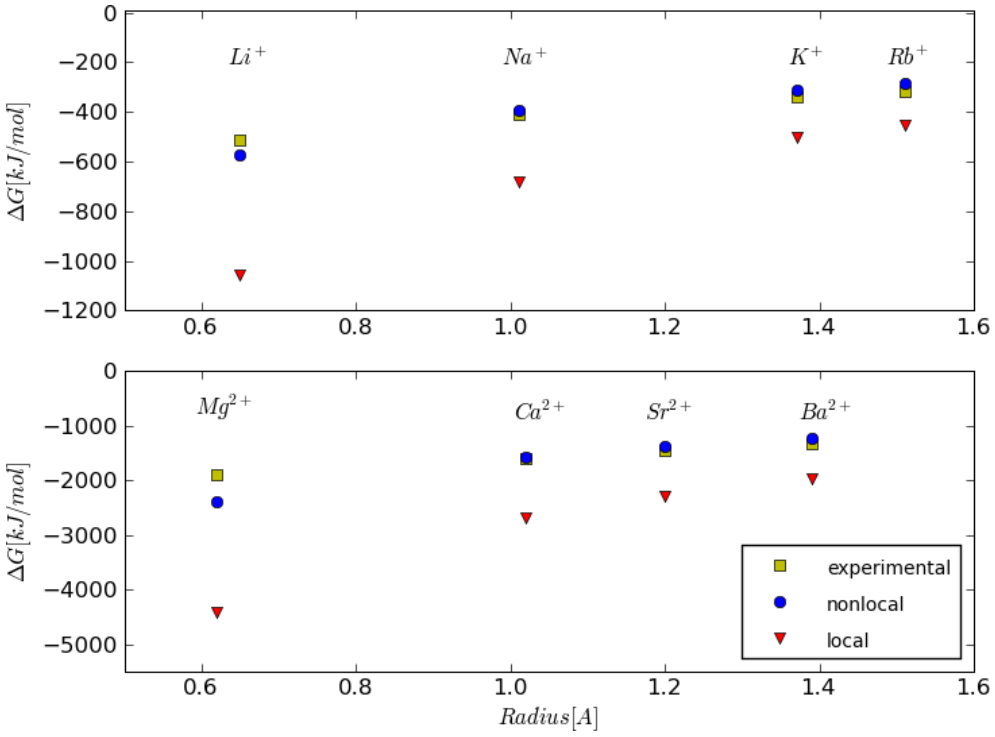


Fig. 3.12: The free energy of solvation of monoatomic ions. The experimental values are taken from [87]; $\kappa = 1/23$ Å^{-1}, $\varepsilon_\Omega = 1\,\varepsilon_0$, $\varepsilon_\Sigma = 78.0\,\varepsilon_0$, $\varepsilon_\infty = 1.8\,\varepsilon_0$.

between 3-7Å. In [52], A. Hildebrandt optimized the correlation length on small amino acid side chains, where he took a range of 2-4Å. In the following we assume λ to be 3.028Å ($\kappa = 1/20$Å^{-1}) unless otherwise mentioned.

As already addressed in Section 3.1.4, the electrostatic energy is only one contribution to the total free enthalpy which describes the solvation process. However, for monoatomic ions with a radius <2Å the nonpolar contributions are negligible [33] in comparison to the electrostatic contributions, because the nonpolar contributions are found to increase with the molecular surface area. Thus, we directly compare the change in electrostatic energy with experimentally measured solvation energies.

Fig. 3.12 illustrates the solvation energies of monoatomic ions for the nonlocal and the local model in comparison with experimental solvation data. The evaluations have been done via Eq. (3.2) [8] with the appropriate potentials.

In Fig. 3.12 we see that the local electrostatic theory cannot sufficiently reproduce the experimental data. The Lorentzian water model highly improves the energy estimation, and thus clearly demonstrates the essential role of spatial dispersion.

Besides the solvation energies, which served as an example here, Kornyshev et al. investigated a number of physical quantities further demonstrating the success of the Lorentzian water model: the hydration energy of ions, the screening effect of ionic fields, the interaction energy of solvated ions, and the solvation energies of the transfer of ions between different solvents, see for example [33–35] and references (25-28) and (36-28) in [124]. However, all these previous studies focus on simple, spherical systems. These systems are highly affected by the nonlocal reaction of the solvent and suggest an important influence for biomolecules, which are naturally solvated in water, as well.

[8] In Section 3.2.2.1 (Theorem 3.3.2), we have shown that this formula is valid for the Lorentzian response.

Chapter 4

Reformulation of the integro-differential equation

In this chapter, we aim at finding a way to apply and solve the nonlocal equations for molecules of arbitrary shape. The approaches discussed in Section 3.3.3, the Fourier transformation for an accumulation of point charges and the integration of the correlation field $\boldsymbol{F}$ in the case of spherical symmetry, are not applicable for arbitrary geometries. The reason is the complexity of the nonlocal electrostatic equations stated in Theorem 3.3.1[1]:

$$\boldsymbol{F}_\Sigma = -\frac{\kappa^2}{4\pi}\int\limits_\Sigma d\boldsymbol{r}'\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\boldsymbol{\nabla}_{r'}\phi_\Sigma(\boldsymbol{r}')\,, \qquad \text{in } \Sigma \tag{4.1a}$$

$$\varepsilon_\Omega\triangle\phi_\Omega = -\rho\,, \qquad \text{in } \Omega \tag{4.1b}$$

$$\varepsilon_\infty\triangle\phi_\Sigma - (\varepsilon_\Sigma-\varepsilon_\infty)\,\boldsymbol{\nabla}\cdot\boldsymbol{F}_\Sigma = 0\,, \qquad \text{in } \Sigma \tag{4.1c}$$

$$\varepsilon_\Omega\partial_n\phi_\Omega - (\varepsilon_\infty\partial_n\phi_\Sigma - (\varepsilon_\Sigma-\varepsilon_\infty)\,(\boldsymbol{n}\cdot\boldsymbol{F}_\Sigma)) = 0\,, \qquad \text{on } \Gamma \tag{4.1d}$$

$$\phi_\Omega - \phi_\Sigma = 0\,, \qquad \text{on } \Gamma \tag{4.1e}$$

We have to solve for a transmission problem defined by an integral, Eq. (4.1a) and by a differential system (4.1)(b-e). To make an application of nonlocal electrostatics to biomolecules feasible, we need an approach which copes with arbitrarily shaped transmission regions Γ. And even more important, we have to overcome the interplay of differential and integral expressions, i.e., we have to cast the system into a purely partial differential equation system (PDES).

Thus, the aim of this chapter is to substitute Eq. (4.1a) by a differential system characterized by a linear operator $\mathcal{L}$:

$$\boxed{\boldsymbol{F}_\Sigma = \frac{\kappa^2}{4\pi}\int\limits_\Sigma d\boldsymbol{r}'\frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\boldsymbol{E}_\Sigma(\boldsymbol{r}'), \quad \text{in } \Sigma} \quad \Rightarrow \quad \boxed{\begin{array}{ll}\mathcal{L}\,\boldsymbol{F}_\Sigma = \,?\,, & \text{in}\,\Sigma \\ +\,\text{boundary values} & \text{on}\,\Gamma\end{array}}$$

In order to find an appropriate differential system which substitutes Eq. (4.1a), we proceed as follows: first, a motivation is given in Section 4.1, where we introduce the differential operator, $\mathcal{L}_\kappa = (\triangle - \kappa^2)$, the so called Yukawa operator. A first analysis of its fundamental solution shows promise for finding a differential formulation. It clarifies further that the differential formulation is unique only when the boundary values are specified.

In Section 4.2 we therefore introduce the basic ideas to uniquely represent an integral formulation as a system of differential equations. To this end, we concentrate on general transmission and

[1] As we assume the molecules not to have a surface charge, we set $\sigma = 0$.

boundary value problems of the Yukawa operator $\mathcal{L}_\kappa$. For both, the transmission as well as the boundary value problem, we deduce unique integral representation formulas.

With these considerations, we start in Section 4.3 with the purely differential formulation of the nonlocal equations of biomolecules immersed in water. In Section 4.3.1, we reinterpret the Lorentz model given in Theorem 3.3.1 as a Newton potential approach. Further in Section 4.3.2, we propose a second model, the so called Dirichlet model, to account for the molecular surface as a disturbing factor of the nonlocal water correlations.

4.1 The Yukawa operator

Our aim is to transform the integral equation of the correlation field $\boldsymbol{F}$ into a system of differential equations. The idea of such a transformation is to search for a linear operator $\mathcal{L}$ whose solution *is* the integral in Eq. (4.1a). An example for such an equivalence is the so called *Yukawa* operator,

$$\mathcal{L}_\kappa := (\triangle - \kappa^2), \quad \kappa \in \mathbb{R}^+$$

together with its Newton potential [2]. In order to show this, we state the fundamental solution of the Yukawa operator:

Theorem 4.1.1 *The fundamental solution of the Yukawa operator: A solution to the following differential equation defined in* $\mathbb{R}^3$

$$\mathcal{L}_\kappa G_{\mathcal{L}_\kappa}(\boldsymbol{r}) = -\delta(\boldsymbol{r}), \quad \boldsymbol{r} \in \mathbb{R}^3 \tag{4.2}$$

is given by

$$\begin{aligned} G_{\mathcal{L}_\kappa} : \quad \mathbb{R}^3\backslash\{0\} &\mapsto \mathbb{R} \\ \boldsymbol{r} &\mapsto \frac{1}{4\pi}\frac{e^{-\kappa|\boldsymbol{r}|}}{|\boldsymbol{r}|}. \end{aligned} \tag{4.3}$$

$G_{\mathcal{L}_\kappa}$ is called the fundamental solution of the Yukawa operator $\mathcal{L}_\kappa$. $G_{\mathcal{L}_\kappa}$ is the unique, physical solution to Eq. (4.2) when the radiation condition has to be fulfilled.
Proof: see [121].

With $G_{\mathcal{L}_\kappa}$, the Newton potential f of the Yukawa operator with source term ρ is defined by

$$f(\boldsymbol{r}) := (G_{\mathcal{L}_\kappa} * \rho)(\boldsymbol{r}) := \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r}'), \quad \boldsymbol{r} \in \mathbb{R}^3, \tag{4.4}$$

as

$$\mathcal{L}_\kappa f(\boldsymbol{r}) = \mathcal{L}_\kappa((G_{\mathcal{L}_\kappa} * \rho)(\boldsymbol{r})) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, \underbrace{\mathcal{L}_\kappa G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')}_{-\delta(\boldsymbol{r}-\boldsymbol{r}')}\rho(\boldsymbol{r}') = -\rho(\boldsymbol{r}), \quad \boldsymbol{r} \in \mathbb{R}^3. \tag{4.5}$$

This means that the Newton potential, Eq. (4.4) is the unique solution of the *globally valid* inhomogeneous Yukawa equation (4.2), which in addition fulfills the radiation condition.

If we recall once again Eq. (4.1a)

$$\boldsymbol{F}_\Sigma(\boldsymbol{r}) = \int\limits_\Sigma d\boldsymbol{r}'\, \frac{e^{-\kappa|\boldsymbol{r}-\boldsymbol{r}'|}}{4\pi\,|\boldsymbol{r}-\boldsymbol{r}'|}(-\kappa^2\,\boldsymbol{\nabla}_{r'}\phi(\boldsymbol{r}')) = \int\limits_\Sigma d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\,\underbrace{(-\kappa^2\,\boldsymbol{\nabla}_{r'}\phi(\boldsymbol{r}'))}_{=:\boldsymbol{\rho}}, \quad \boldsymbol{r} \in \Sigma,$$

[2] The nomenclature has been introduced in Section 2.1.1 for the Laplace operator.

we observe a similarity to Eq. (4.4), which lets us hope that we already found a way to reformulate the integral equation: an application of the Yukawa operator $\mathcal{L}_\kappa$ on $\boldsymbol{F}_\Sigma$ reveals the partial differential equation

$$\mathcal{L}_\kappa \boldsymbol{F}_\Sigma = \kappa^2 \boldsymbol{\nabla} \phi_\Sigma, \quad \boldsymbol{r} \in \Sigma\,. \tag{4.6}$$

In other words, a solution to Eq. (4.6) is

$$\boldsymbol{F}_\Sigma(\boldsymbol{r}) = -\kappa^2 \int\limits_\Sigma d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\, \boldsymbol{\nabla}_{r'} \phi_\Sigma(\boldsymbol{r}'), \quad \boldsymbol{r} \in \Sigma\,.$$

Remark 4.1 Assume $f_0 \in \mathrm{Kern}(\mathcal{L}_\kappa)$, which means $\mathcal{L}_\kappa f_0 = 0$ in Σ. Then, $(\boldsymbol{F}_\Sigma + f_0)$ is also a solution to Eq. (4.6). If $\Sigma = \mathbb{R}^3$, the radiation condition allows only for $f_0 \equiv 0$ in $\mathbb{R}^3$ and we end up in the vectorial analog of Eq. (4.4).

With Remark 4.1, we know that in order to obtain the unique solution, we have to *specify* the behavior of $\boldsymbol{F}_\Sigma$ on Γ, i.e., we have to specify the boundary values. The focus of the following section is the representation of the behavior on Γ for general differential problems of the Yukawa operator.

$$\boxed{\boldsymbol{F}_\Sigma = -\kappa^2 \int\limits_\Sigma d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\, \boldsymbol{\nabla}_{r'} \phi_\Sigma(\boldsymbol{r}'), \quad \text{in } \Sigma} \quad \Rightarrow \quad \boxed{\begin{array}{ll} \mathcal{L}_\kappa\, \boldsymbol{F}_\Sigma = \kappa^2 \boldsymbol{\nabla} \phi_\Sigma\,, & \text{in } \Sigma \\ + \text{boundary values} & \text{on } \Gamma \end{array}}$$

4.2 Boundary and transmission problem of the Yukawa operator

The study to find *boundary integral representations* is closely connected with the task to find solutions of an elliptic PDES within a variational formulation. Since the mathematical theory goes beyond the focus of this work, we only motivate and state the important results. For a more general and detailed analysis we refer to Steinbach's textbook [121] and the excellent book series written by R. Dautray and J. L. Lions [28](vol. 1-6).

The basis of an integral representation is Green's second identity [76, 121]: assume two scalar functions u and v, both twice continuously differentiable in $\Omega \subset \mathbb{R}^3$, then, it holds

$$\begin{aligned} \int\limits_\Omega d\boldsymbol{r}(u \triangle v - v \triangle u) &= \int\limits_\Gamma d\Gamma_r\, \boldsymbol{n} \cdot (u \boldsymbol{\nabla} v - v \boldsymbol{\nabla} u) = \int\limits_\Gamma d\Gamma_r\, (u\,(\boldsymbol{n}\cdot\boldsymbol{\nabla})v - v\,(\boldsymbol{n}\cdot\boldsymbol{\nabla})u) \\ &= \int\limits_\Gamma d\Gamma_r\, (\gamma_0^\Omega u\, \gamma_1^\Omega v - \gamma_0^\Omega v\, \gamma_1^\Omega u)\,, \end{aligned} \tag{4.7}$$

where we introduce the trace operators γ_0 and γ_1 in Definition 4.1 on the next page. It is easy to show that an equation analogous to Eq. (4.7) exists for the Yukawa operator $\mathcal{L}_\kappa$:

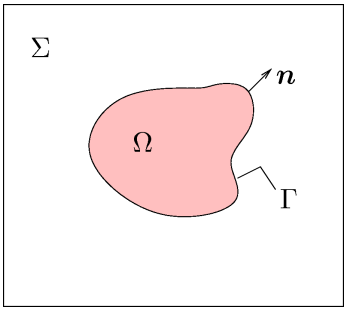
$$\int\limits_\Omega d\boldsymbol{r}(u\, \mathcal{L}_\kappa v - v\, \mathcal{L}_\kappa u) = \int\limits_\Gamma d\Gamma_r\, (\gamma_0^\Omega u\, \gamma_1^\Omega v - \gamma_0^\Omega v\, \gamma_1^\Omega u)\,, \tag{4.8}$$

Definition 4.1 The trace operators:

Internal Dirichlet trace on Γ: $\gamma_0^\Omega f := \lim_{\Omega \ni \boldsymbol{r} \to \Gamma} f(\boldsymbol{r})$

Internal Neumann trace on Γ: $\gamma_1^\Omega f := \lim_{\Omega \ni \boldsymbol{r} \to \Gamma} (\boldsymbol{n} \cdot \boldsymbol{\nabla}) f(\boldsymbol{r})$

As illustrated on the right, the normal vector on Γ points into Σ, i.e., it is the outer normal of Ω.



Now, we consider the following differential equation in the confined region Ω

$$\mathcal{L}_\kappa f = -\rho_\Omega, \quad \text{in } \Omega\,. \tag{4.9}$$

Green's second identity applied on the solution f of Eq. (4.9) and the fundamental solution of the Yukawa operator $\mathcal{L}_\kappa$, i.e., replacing u with f and v with $G_{\mathcal{L}_\kappa}$ in Eq. (4.8), reads

$$\begin{aligned}
&\int_\Omega d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\,(\mathcal{L}_\kappa f(\boldsymbol{r}')) \\
&\quad = \int_\Omega d\boldsymbol{r}'\,(\mathcal{L}_\kappa G_{\mathrm{Y}}(\boldsymbol{r}-\boldsymbol{r}'))\, f(\boldsymbol{r}') + \int_\Gamma d\Gamma_{r'}\,\left[\gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\;\gamma_1^\Omega f(\boldsymbol{r}') - \gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\;\gamma_0^\Omega f(\boldsymbol{r}')\right] \\
&\quad = -\int_\Omega d\boldsymbol{r}'\,\delta(\boldsymbol{r}-\boldsymbol{r}')\, f(\boldsymbol{r}') + \int_\Gamma d\Gamma_{r'}\,\left[\gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\;\gamma_1^\Omega f(\boldsymbol{r}') - \gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\;\gamma_0^\Omega f(\boldsymbol{r}')\right]\,,
\end{aligned}$$

This reveals the so called *representation formula* of the solution of Eq. (4.9)

$$\begin{aligned}
f(\boldsymbol{r}) \;=\; & \int_\Omega d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\rho_\Omega(\boldsymbol{r}') + \int_\Gamma d\Gamma_{r'}\,\gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\;\gamma_1^\Omega f(\boldsymbol{r}') \\
& - \int_\Gamma d\Gamma_{r'}\,\gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\,\gamma_0^\Omega f(\boldsymbol{r}'), \qquad \text{in } \Omega\,.
\end{aligned} \tag{4.10}$$

Important remark 4.1 In order to calculate the solution f at position $\boldsymbol{r} \in \Omega$, the integral formula Eq. (4.10) requires both traces, $\gamma_0^\Omega f$ and $\gamma_1^\Omega f$. From the physical point of view, it is clear that either the Dirichlet or the Neumann trace is necessary to uniquely define the physical problem [61, 97]. The remaining trace is determined by the physical setting.

In consideration of Remark 4.1, we now formulate the solution of the Dirichlet and the Neumann boundary value problem, respectively.

Theorem 4.2.1 *Boundary value problems of the Yukawa operator:*
The function

$$\begin{aligned}
f(\boldsymbol{r}) \;=\; & \int_\Omega d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\rho_\Omega(\boldsymbol{r}') + \int_\Gamma d\Gamma_{r'}\,\gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\;\gamma_1^\Omega f(\boldsymbol{r}') \\
& - \int_\Gamma d\Gamma_{r'}\,\gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\,\gamma_0^\Omega f(\boldsymbol{r}'), \qquad \text{in } \Omega\,,
\end{aligned} \tag{4.11}$$

is the solution to the following boundary value problems:

1. *Eq.* (4.11) *is the solution to the Dirichlet boundary value problem*

$$\left.\begin{array}{ll} \mathcal{L}_\kappa f = -\rho_\Omega, & \text{in } \Omega \\ \gamma_0^\Omega f = g_0^\Omega, & \text{on } \Gamma \end{array}\right\} \tag{4.12}$$

   where the Neumann data $\gamma_1^\Omega f$ *in Eq.* (4.11) *is implicitly given by system* (4.12).

2. *Eq.* (4.11) *is the solution to the Neumann boundary value problem*

$$\left.\begin{array}{ll} \mathcal{L}_\kappa f = -\rho_\Omega, & \text{in } \Omega \\ \gamma_1^\Omega f = g_1^\Omega, & \text{on } \Gamma \end{array}\right\} \tag{4.13}$$

   where the Dirichlet data $\gamma_0^\Omega f$ *in Eq.* (4.11) *is implicitly given by system* (4.13).

An analogous representation to Eq. (4.10) holds in Σ. With the convention that the surface normal points into Σ as illustrated in Definition 4.1, we obtain

$$\begin{aligned} f(\boldsymbol{r}) \;=\; & \int_\Sigma d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\rho_\Sigma(\boldsymbol{r}') - \int_\Gamma d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; \gamma_1^\Sigma f(\boldsymbol{r}') \\ & + \int_\Gamma d\Gamma_{r'}\, \gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\, \gamma_0^\Sigma f(\boldsymbol{r}'), \qquad \text{in } \Sigma\,. \end{aligned} \tag{4.14}$$

We conclude that a general representation formula for the solution f of the following transmission problem

$$\left.\begin{array}{lcll} \mathcal{L}_\kappa f &=& -\rho_\Omega, & \text{in } \Omega \\ \mathcal{L}_\kappa f &=& -\rho_\Sigma, & \text{in } \Sigma \\ [\gamma_0 f] &=& \gamma_0^\Omega f - \gamma_0^\Sigma f & \text{on } \Gamma \\ [\gamma_1 f] &=& \gamma_1^\Omega f - \gamma_1^\Sigma f & \text{on } \Gamma \end{array}\right\} \tag{4.15}$$

can be derived from the representation formulas in Ω (Eq. (4.10)), and Σ (Eq. (4.14)):

$$\begin{aligned} f(\boldsymbol{r}) \;=\; & \underbrace{\int_{\mathbb{R}^3} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r}')}_{\text{Newton potential}} + \underbrace{\int_\Gamma d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; [\gamma_1 f]}_{\text{Single layer potential}} \\ & \underbrace{- \int_\Gamma d\Gamma_{r'}\, \gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; [\gamma_0 f]}_{\text{Double layer potential}}, \qquad \boldsymbol{r} \in \Omega \cup \Sigma \end{aligned} \tag{4.16}$$

with

$$\rho = \left\{ \begin{array}{ll} \rho_\Omega, & \text{in } \Omega \\ \rho_\Sigma, & \text{in } \Sigma\,. \end{array}\right. \tag{4.17}$$

Important remark 4.2

- The representation formulas (4.10), (4.14) and (4.16) tell us that we can express the solution of the boundary and transmission value problems of the Yukawa operator $\mathcal{L}_\kappa$, (4.12), (4.13) and (4.15), by two boundary integrals and one volume integral. The boundary integrals must appear in the representation formulas in order to incorporate the boundary conditions.

- The derivation is valid for the Yukawa operator with $\kappa \geq 0$ and therefore it holds for the Laplace operator $\triangle$ as well. This is consistent with the fundamental solution $G_{\triangle}(\boldsymbol{r}-\boldsymbol{r}')$ defined in Section 2.1.1.
- The volume and boundary integrals that appear in the general representation formulas, have different mathematical features and different physical meanings [28, 29]. Exemplarily, we denoted the different integrals in Eq. (4.16) and give here their mathematical definitions [28, 121]. A definition and a short motivation of the used function spaces is given in Appendix 10.1:

  Definition 4.2 Newton potential: Let $\rho \in H^{-1}(\mathbb{R}^3)$ be a given density function. Then,

  $$\begin{aligned} \tilde{N}_0 &: H^{-1}(\mathbb{R}^3) \mapsto H^1(\mathbb{R}^3) \\ (\tilde{N}_0\rho)(\boldsymbol{r}) &:= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r}') = (G_{\mathcal{L}_\kappa} * \rho)(\boldsymbol{r}), \quad \boldsymbol{r} \in \mathbb{R}^3\,, \end{aligned}$$

  is the Newton potential of the Yukawa operator $\mathcal{L}_\kappa$.

  Definition 4.3 Single layer potentials: Let $w \in H^{-1/2}(\Gamma)$ be a given density function. Then,

  $$\begin{aligned} \tilde{V}^{\mathrm{Y}} &: H^{-1/2}(\Gamma) \mapsto H^1(\Omega \cup \Sigma) \\ (\tilde{V}^{\mathrm{Y}} w)(\boldsymbol{r}) &:= \int\limits_{\Gamma} d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')w(\boldsymbol{r}'), \quad \text{in } \Omega \cup \Sigma \end{aligned}$$

  is the Single layer potential of the Yukawa operator $\mathcal{L}_\kappa$ and

  $$\begin{aligned} \tilde{V}^{\mathrm{L}} &: H^{-1/2}(\Gamma) \mapsto H^1(\Omega \cup \Sigma) \\ (\tilde{V}^{\mathrm{L}} w)(\boldsymbol{r}) &:= \int\limits_{\Gamma} d\Gamma_{r'}\, \gamma_{0,r'} G_{\triangle}(\boldsymbol{r}-\boldsymbol{r}')w(\boldsymbol{r}'), \quad \text{in } \Omega \cup \Sigma \end{aligned}$$

  is the Single layer potential of the Laplace operator $\triangle$.

  Definition 4.4 Double layer potentials: Let $v \in H^{1/2}(\Gamma)$ be a given density function. Then,

  $$\begin{aligned} W^{\mathrm{Y}} &: H^{1/2}(\Gamma) \mapsto H^1(\Omega \cup \Sigma) \\ (W^{\mathrm{Y}} v)(\boldsymbol{r}) &:= \int\limits_{\Gamma} d\Gamma_{r'}\, \gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')v(\boldsymbol{r}'), \quad \text{in } \Omega \cup \Sigma \end{aligned}$$

  defines the Double layer potential of the Yukawa operator $\mathcal{L}_\kappa$ and

  $$\begin{aligned} W^{\mathrm{L}} &: H^{1/2}(\Gamma) \mapsto H^1(\Omega \cup \Sigma) \\ (W^{\mathrm{L}} v)(\boldsymbol{r}) &:= \int\limits_{\Gamma} d\Gamma_{r'}\, \gamma_{1,r'} G_{\triangle}(\boldsymbol{r}-\boldsymbol{r}')v(\boldsymbol{r}'), \quad \text{in } \Omega \cup \Sigma \end{aligned}$$

  is the Double layer potential of the Laplace operator $\triangle$.

4.3 Generalization of the correlation field in biomolecular applications

The outcome of the previous considerations is the following: we found a general integral representation of the boundary value problem of the Yukawa operator in Σ:

$$\boxed{\begin{aligned} f_\Sigma(\boldsymbol{r}) &= \int\limits_{\Sigma} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\rho_\Sigma(\boldsymbol{r}') \\ &\quad - \int\limits_{\Gamma} d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\ \gamma_1^\Sigma f(\boldsymbol{r}') \\ &\quad + \int\limits_{\Gamma} d\Gamma_{r'}\, \gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\, \gamma_0^\Sigma f(\boldsymbol{r}'), \text{ in } \Sigma \end{aligned}} \quad \Leftrightarrow \quad \boxed{\begin{aligned} \mathcal{L}_\kappa f_\Sigma &= -\rho_\Sigma, \quad \text{in } \Sigma \\ \underbrace{\gamma_0^\Sigma f_\Sigma = g_0^\Sigma}_{\text{Dirichlet problem}} &, \quad \text{on } \Gamma \\ \underbrace{\gamma_1^\Sigma f_\Sigma = g_1^\Sigma}_{\text{Neumann problem}} &, \quad \text{on } \Gamma \end{aligned}} \quad (*)$$

Using Eq. (4.17), we found a general integral representation of the transmission problem of the Yukawa operator in $\mathbb{R}^3$:

$$\boxed{\begin{aligned} f(\boldsymbol{r}) \;=\; & \int_{\mathbb{R}^3} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\rho(\boldsymbol{r}') \\ & + \int_{\Gamma} d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; [\gamma_1 f] \\ & - \int_{\Gamma} d\Gamma_{r'}\, \gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; [\gamma_0 f], \text{ in } \Omega\cup\Sigma \end{aligned}} \quad \Leftrightarrow \quad \boxed{\begin{aligned} \mathcal{L}_\kappa f_\Omega &= -\rho_\Omega, \quad \text{in } \Omega \\ \mathcal{L}_\kappa f_\Sigma &= -\rho_\Sigma, \quad \text{in } \Sigma \\ [\gamma_0 f] &= [\gamma_0 g], \quad \text{on } \Gamma \\ [\gamma_1 f] &= [\gamma_1 g], \quad \text{on } \Gamma \end{aligned}} \qquad (**)$$

With this knowledge, we now try to find an appropriate system of differential equations to replace the integral representation of the correlation field $\boldsymbol{F}$,

$$\boldsymbol{F}_\Sigma(\boldsymbol{r}) = -\kappa^2 \int_\Sigma d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\, \boldsymbol{\nabla}_{r'}\phi_\Sigma(\boldsymbol{r}'), \quad \boxed{\boldsymbol{r}\in\Sigma} \tag{4.18}$$

in Theorem 3.3.1, i.e., in the electrostatic material equations of the biomolecular system.

Here, we emphasize that Eq. (4.18) is restricted to Σ – without any boundary integrals over Γ that further determine the correlation field $\boldsymbol{F}_\Sigma$ on Γ. In contrast, both, the boundary value problem given in $(*)$ as well as the transmission problem given in $(**)$, are only well defined for given boundary and jump conditions, respectively.

In Section 4.3.1 we find that the correlation field $\boldsymbol{F}_\Sigma$ as given in Eq. (4.18) can be identified from a transmission problem with vanishing jumps on Γ, i.e., it is a Newton potential:

$$\boldsymbol{F}_{\mathrm{N}}(\boldsymbol{r}) \;=\; -\kappa^2 \int_\Sigma d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{\nabla}_{r'}\phi_\Sigma(\boldsymbol{r}'), \quad \boxed{\boldsymbol{r}\in\mathbb{R}^3}$$

Further, in Section 4.3.1 we state the purely differential system equivalent to Theorem 3.3.1 and solve the differential system for the Born sphere in Section 4.3.1.1. The whole working procedure is shown on the left side of the diagram 4.1.

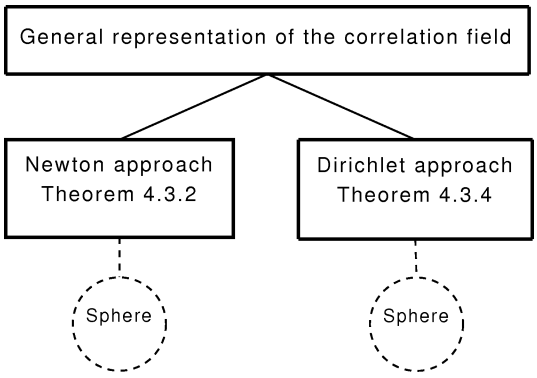


Fig. 4.1: Overview of the two different approaches for the Lorentzian water model in the biomolecular system.

However, the potential of this novel, general representation formula lies in the boundary value approach given in (∗). It bears freedom to define a specific behavior on the surface Γ, which for example can be physically motivated or given by experimental measurements. Specifying the correlation field $\boldsymbol{F}$ on the surface Γ goes beyond the established representation of the correlation field as a pure volume integral.

In Section 4.3.2 we will argue for a Dirichlet approach as shown on the right of the diagram 4.1. In fact, we restrict to a vanishing Dirichlet trace

$$\gamma_0^{\Sigma}\boldsymbol{F}_{\mathrm{D},\Sigma} = \boldsymbol{0}, \quad \text{on } \Gamma\,,$$

which leads to a modification of the "original" definition of $\boldsymbol{F}$:

$$\boldsymbol{F}_{\mathrm{D},\Sigma}(\boldsymbol{r}) = \underbrace{-\kappa^2 \int_{\Sigma} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{\nabla}_{r'}\phi_\Sigma(\boldsymbol{r}')}_{\text{original definition of } \boldsymbol{F}} - \int_{\Gamma} d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; \gamma_1^{\Sigma} F_{\mathrm{D},\Sigma}(\boldsymbol{r}'), \quad \boxed{\boldsymbol{r}\in\Sigma}$$

4.3.1 Newton potential approach

In order to find an appropriate PDES for the correlation field $\boldsymbol{F}_\Sigma$,

$$\boldsymbol{F}_\Sigma(\boldsymbol{r}) = -\kappa^2 \int_{\Sigma} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\, \boldsymbol{\nabla}_{r'}\phi_\Sigma(\boldsymbol{r}'), \quad \text{in } \Sigma\,, \tag{4.19}$$

we start with the representation formula of a transmission problem (∗∗):

$$\begin{aligned} \boldsymbol{F}(\boldsymbol{r}) \;&=\; \int_{\mathbb{R}^3} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{\rho}(\boldsymbol{r}') \\ &\quad + \int_{\Gamma} d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; [\gamma_1 \boldsymbol{F}] - \int_{\Gamma} d\Gamma_{r'}\, \gamma_{1,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; [\gamma_0 \boldsymbol{F}], \text{ in } \mathbb{R}^3 \end{aligned} \tag{4.20}$$

Determining the jump conditions $[\gamma_0\boldsymbol{F}]$, $[\gamma_1\boldsymbol{F}]$ and the source term $\boldsymbol{\rho}$ in a way that

$$\boldsymbol{F}_\Sigma(\boldsymbol{r}) \equiv \boldsymbol{F}(\boldsymbol{r})\big|_\Sigma\,,$$

gives the representation formula and the corresponding PDES that – restricted to Σ – equals the correlation field $\boldsymbol{F}_\Sigma$ given in Eq. (4.19). A comparison of Eq. (4.19) and Eq. (4.20) reveals the following conditions

$$\begin{aligned} [\gamma_0\boldsymbol{F}] &= \boldsymbol{0}\,, \quad \text{on } \Gamma \\ [\gamma_1\boldsymbol{F}] &= \boldsymbol{0}\,, \quad \text{on } \Gamma \end{aligned}$$

$$\boldsymbol{\rho}(\boldsymbol{r}) \;=\; -\kappa^2\chi_\Sigma \boldsymbol{\nabla}\phi_\Sigma(\boldsymbol{r}) \;=\; \begin{cases} \boldsymbol{0}\,, & \text{in } \Omega \\ -\kappa^2\boldsymbol{\nabla}\phi_\Sigma(\boldsymbol{r})\,, & \text{in } \Sigma\,, \end{cases} \tag{4.21}$$

where we used the definition in Eq. (3.20) made on p. 39 of the characteristic function χ_V.

As the jumps of the Dirichlet and the Neumann traces vanish in the representation formula (4.20), the correlation field $\boldsymbol{F}_\Sigma$ can be interpreted as a Newton potential. This is denoted by the index N

in the remainder of this work:

$$\begin{aligned} \boldsymbol{F}_{\mathrm{N}}(\boldsymbol{r}) &= -\kappa^2 \int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\, \chi_\Sigma \boldsymbol{\nabla}_{r'} \phi_\Sigma(\boldsymbol{r}') \\ &= -\kappa^2 \int\limits_{\Sigma} d\boldsymbol{r}'\, G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\, \boldsymbol{\nabla}_{r'} \phi_\Sigma(\boldsymbol{r}')\,, \quad \boldsymbol{r} \in \mathbb{R}^3 \end{aligned} \tag{4.22}$$

We finally obtain Theorem 4.3.1.

Theorem 4.3.1 *Reformulation of the correlation field $\boldsymbol{F}$ in the Newton potential approach: Let $\mathbb{R}^3$ be decomposed into $\Omega \cup \Gamma \cup \Sigma$. The integral equation*

$$\boldsymbol{F}_N(\boldsymbol{r}) = -\kappa^2 \int\limits_{\Sigma} d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\, \boldsymbol{\nabla}_{r'} \phi_\Sigma(\boldsymbol{r}'), \quad \boldsymbol{r} \in \mathbb{R}^3$$

is the unique solution of the transmission problem:

$$\begin{aligned} \mathcal{L}_\kappa \boldsymbol{F}_{N,\Omega} &= \boldsymbol{0}, && \text{in } \Omega \\ \mathcal{L}_\kappa \boldsymbol{F}_{N,\Sigma} &= \kappa^2 \boldsymbol{\nabla} \phi_\Sigma, && \text{in } \Sigma \\ \boldsymbol{F}_{N,\Omega} &= \boldsymbol{F}_{N,\Sigma}, && \text{on } \Gamma \\ (\boldsymbol{n}\cdot\boldsymbol{\nabla})\boldsymbol{F}_{N,\Omega} &= (\boldsymbol{n}\cdot\boldsymbol{\nabla})\boldsymbol{F}_{N,\Sigma}, && \text{on } \Gamma \end{aligned}$$

Now, we can replace the integral representation of the correlation field in the Lorentzian model of the biomolecular system (Theorem 3.3.1, p. 41) and end up with its differential analog, i.e., the Lorentzian model of the biomolecular system in the Newton approach:

Theorem 4.3.2 *The Lorentzian model of the biomolecular system in differential form in the Newton approach:*

$$\begin{aligned} \varepsilon_\Omega \triangle \phi_\Omega &= -\rho, && \text{in } \Omega \\ \varepsilon_\infty \triangle \phi_\Sigma - (\varepsilon_\Sigma - \varepsilon_\infty)\,(\boldsymbol{\nabla}\cdot\boldsymbol{F}_{N,\Sigma}) &= 0, && \text{in } \Sigma \\ \varepsilon_\Omega \partial_n \phi_\Omega &= \varepsilon_\infty \partial_n \phi_\Sigma - (\varepsilon_\Sigma - \varepsilon_\infty)\,(\boldsymbol{n}\cdot\boldsymbol{F}_{N,\Sigma}), && \text{on } \Gamma \\ \phi_\Omega &= \phi_\Sigma, && \text{on } \Gamma \end{aligned}$$

$$\begin{aligned} \mathcal{L}_\kappa \boldsymbol{F}_{N,\Omega} &= \boldsymbol{0}, && \text{in } \Omega \\ \mathcal{L}_\kappa \boldsymbol{F}_{N,\Sigma} &= \kappa^2 \boldsymbol{\nabla} \phi_\Sigma, && \text{in } \Sigma \\ \boldsymbol{F}_{N,\Omega} - \boldsymbol{F}_{N,\Sigma} &= \boldsymbol{0}, && \text{on } \Gamma \\ (\boldsymbol{n}\cdot\boldsymbol{\nabla})\boldsymbol{F}_{N,\Omega} - (\boldsymbol{n}\cdot\boldsymbol{\nabla})\boldsymbol{F}_{N,\Sigma} &= \boldsymbol{0}, && \text{on } \Gamma \end{aligned}$$

In the remainder of this work, this vector model for nonlocal electrostatics of the biomolecular system is called the *Newton Vector Model* (NVM).

The second block in Theorem 4.3.1 describes the nonlocal correlation field in Σ and, although the nonlocal response theory has been applied in Σ, there is a connection to Ω. From the mathematical point of view, this influence originates from the requirement on *global* validity of $\boldsymbol{F}_{\mathrm{N}}$.

From a physical point of view, the equation in Ω is *a priori* unclear, because the nonlocal response, which is defined by $\boldsymbol{F}_{\mathrm{N},\Sigma}$ takes place in Σ only. However, the influence of $\boldsymbol{F}_{\mathrm{N},\Omega}$ on $\boldsymbol{F}_{\mathrm{N},\Sigma}$ can be interpreted in terms of the nonlocal correlation, which are in this model rather mediated than disturbed by Ω. Having this in mind, the solution $\boldsymbol{F}_{\mathrm{N},\Omega}$ itself has no physical meaning,

but it implies a smooth decrease of the nonlocal correlations towards Ω. This co-determines the transmission condition and therefore influences the nonlocal polarization in Σ.

The mathematical meaning of the original correlation field $\boldsymbol{F}$ as Newton potential $\boldsymbol{F}_{\mathrm{N}}$ was first observed by C. Fasel and S. Rjasanow [39, 40]. Their analysis, however, aimed at solving the corresponding differential equations. The physical meaning of the correlation field $\boldsymbol{F}$ was not investigated. The novel formulation reveals a physical insight: we interpret the Newton potential approach as the conceptually simplest treatment of biomolecules immersed in water, because the nonlocal dielectric response does not experience any variations due to the dielectric boundary.

4.3.1.1 The Born model revisited

In Section 3.3.3, the integro-differential system for the Born model was solved by an explicit integration. Now, we want to use the purely differential formalism to find the solution of the Born model. Because of the spherical symmetry, all fields can be expressed by gradients of potentials. We make the ansatz

$$\boldsymbol{F}_{\mathrm{N}} := -\boldsymbol{\nabla} F_{\mathrm{N}}\,, \quad \text{in } \mathbb{R}^3\,,$$

and solve the partial differential system given in Theorem 4.3.2 for a point charge, $\rho = q\,\delta(\boldsymbol{r})$ in the center of a sphere with radius a located at the origin of the coordinate system. The system given in Theorem 4.3.2 turns into:

$$\left.\begin{aligned}
\varepsilon_\Omega \triangle \phi_\Omega &= -q\,\delta(r), && r<a\\
\triangle[\varepsilon_\infty \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty) F_{\mathrm{N},\Sigma}] &= 0, && r>a\\
\phi_\Omega &= \phi_\Sigma, && r=a\\
\varepsilon_\Omega \partial_n \phi_\Omega &= \varepsilon_\infty \partial_n \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty)\,\partial_n F_{\mathrm{N},\Sigma}, && r=a\\
\mathcal{L}_\kappa F_{\mathrm{N},\Omega} &= 0, && r<a\\
\mathcal{L}_\kappa F_{\mathrm{N},\Sigma} &= -\kappa^2 \phi_\Sigma, && r>a\\
\partial_n F_{\mathrm{N},\Omega} &= \partial_n F_{\mathrm{N},\Sigma}, && r=a\\
\partial_n^2 F_{\mathrm{N},\Omega} &= \partial_n^2 F_{\mathrm{N},\Sigma}, && r=a
\end{aligned}\right\} \tag{4.23}$$

The analytical solution of system (4.23) is

$$\begin{aligned}
\phi_\Omega &= \frac{1}{4\pi}\left[\frac{1}{\varepsilon_\Omega\, r} + A\right]\\
\phi_\Sigma &= \frac{1}{4\pi r}\left[\frac{1}{\varepsilon_\Sigma} + B\, e^{-s\,(r-a)}\right]\\
F_{\mathrm{N},\Omega} &= \frac{1}{4\pi r}\, B_2(B)\, \sinh(\kappa\, r)\\
F_{\mathrm{N},\Sigma} &= \frac{1}{4\pi r}\left[\frac{1}{\varepsilon_\Sigma} - B\, e^{-s\,(r-a)}\frac{\varepsilon_\infty}{\varepsilon_\Sigma - \varepsilon_\infty}\right],
\end{aligned}$$

where s, B and $B_2(B)$ are defined by

$$\begin{aligned}
s &= \tfrac{1}{\lambda} = \sqrt{\tfrac{\varepsilon_\Sigma}{\varepsilon_\infty}}\kappa\\
A &= \left(\tfrac{1}{\varepsilon_\Sigma} - \tfrac{1}{\varepsilon_\Omega} + B\right)\tfrac{1}{a}\\
B &= \left(\tfrac{1}{\varepsilon_\infty} - \tfrac{1}{\varepsilon_\Sigma}\right)\frac{1}{1 + s\,a + \frac{\varepsilon_\Sigma}{\varepsilon_\infty}\left(\frac{\cosh(\kappa\,a)\,\kappa\,a}{\sinh(\kappa\,a)} - 1\right)}\\
B_2(B) &= B\,\tfrac{1}{\sinh(\kappa a)}\,\tfrac{\varepsilon_\Sigma}{\varepsilon_\infty - \varepsilon_\Sigma} \quad .
\end{aligned}$$

As expected, the solution of the system (4.23) is the same as the one derived in Section 3.3.3.2. Solving the purely differential system is elegant and simple compared to the non-trivial integrations.

Important remark 4.3 When comparing the general solution of F_Σ, Eq. (3.31) on p. 47, with the one derived here, we see that Eq. (3.31) contains a term, which lies in the kernel of the Yukawa operator $\mathcal{L}_\kappa$. This term vanishes only if we demand that F_Σ is subject to the electrostatic material equations. In the case of solving the differential system, this contribution directly vanishes, because the material equations are simultaneously fulfilled:

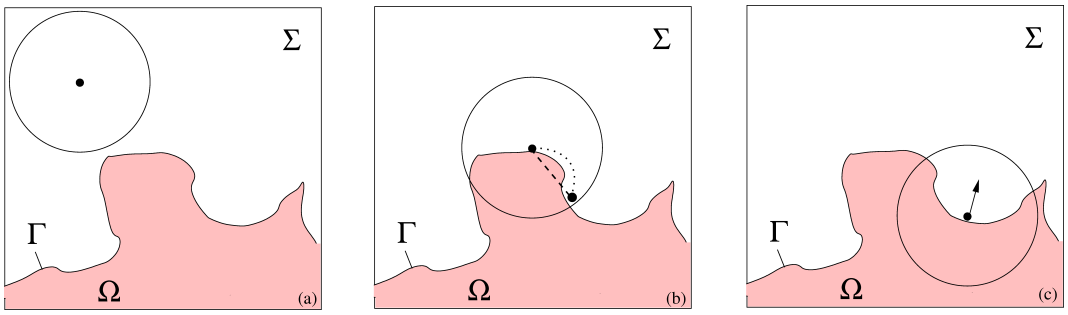
$$\left\{ \begin{array}{rcll} \triangle[\varepsilon_\infty \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty) F_{\mathrm{N},\Sigma}] & = & 0, & r > a \\ \mathcal{L}_\kappa F_{\mathrm{N},\Sigma} + \kappa^2 \phi_\Sigma & = & 0, & r > a \end{array} \right\} \quad \Leftrightarrow \quad \left\{ \triangle \left[(\triangle - \frac{1}{\lambda^2}) F_{\mathrm{N},\Sigma} \right] = 0, \quad r > a \right\}$$

The above equivalence proves that F_Σ does not has a contribution, which lies in $\mathrm{Kern}(\mathcal{L}_\kappa)$, but rather two contributions, one lies in $\mathrm{Kern}(\triangle)$, the other lies in $\mathrm{Kern}(\mathcal{L}_{1/\lambda})$. This means that - at least for the spherically symmetric case - a response on the length scale $\sim 1/\kappa$ is not driven by the system.

Remark 4.2 A reduction of the six transmission conditions to two scalar equations for the correlation field $\boldsymbol{F}_\mathrm{N}$ as done in system (4.23) is possible without further restrictions only in the case of spherical symmetry. Here, the normal direction on the surface Γ coincides with the spatial dependence of all electrostatic fields.

4.3.2 Dirichlet problem

The Lorentzian model of water characterizes a homogeneous situation, i.e., it does not model a disruptive element like a biomolecule. To consider the effect of the biomolecule in the Maxwell equations requires a physical intuition of the way the nonlocal correlations "behave" close to the biomolecule and therefore determine the boundary values of $\boldsymbol{F}_\Sigma$.



Let us recall Fig. 3.9 shown once again above: we sketch three different positions of the correlation sphere (CS) of a single water molecule: in the figures (b-c) the CS penetrates Ω. The nonlocal network, which is mediated in Σ, is disturbed because of the dielectric boundary. When the CS approaches Ω, the water network "feels" the change in the dielectric response. The local formation of the network highly depends on the nature of the amino acids exposed to the surface: for instance, in nonpolar molecular regions we expect the hydrogen bonds to turn away from the surface, whereas in polar regions specific hydrogen bonds of water and amino acids of the protein can often be built. Thus, it is difficult to predict the behavior of the correlation field on Γ without taking into account the local protein composition.

One estimate of the dielectric response on the molecular surface is given by the assumption that the water molecules on the surface respond only with their internal, electronic response ε_∞ [52]. It corresponds to the fact that the molecule binds a first layer of water molecules very tightly in a small region around the surface. This is due to the fact, that the electrostatic field variations are strongest on the surface and thus the orientational polarizability is weakened. Furthermore, the loss in hydrogen bonds in direction Ω, i.e., a loss in entropic energy, can increase the strength of the

hydrogen bonds in normal direction away from the surface, as the molecules keep their remaining hydrogen bonds [112]. This motivates to discuss a model where the correlation field $\boldsymbol{F}_\Sigma$ vanishes on Γ, and thus we introduce the Dirichlet approach with vanishing Dirichlet trace:

$$\gamma_0^\Sigma \boldsymbol{F}_{\mathrm{D},\Sigma} = \lim_{\Sigma \ni \boldsymbol{r} \to \Gamma} \boldsymbol{F}_{\mathrm{D},\Sigma}(\boldsymbol{r}) = \boldsymbol{0}\,, \quad \text{on } \Gamma$$

From a physical point of view, this means that the dielectric response mediated by ε starts with the internal, electronic dielectric function. Moving away from Γ, the network is gradually built up - mainly in surface normal direction, because in this direction the water molecules are less affected by the dielectric boundary Γ.

In Section 3.3.2.1, we learned that a pure electronic response coincides with a frozen water network and that this states the contrary case to the local model of macroscopic response. Having this in mind, a vanishing Dirichlet boundary condition is one possible limiting case for the behavior of the water network on the molecular surface.

Inserting the vanishing Dirichlet trace $\gamma_0^\Sigma \boldsymbol{F}_{\mathrm{D},\Sigma}$ into the general representation formula for the unbounded domain Σ, Eq. (4.14), yields

$$\boldsymbol{F}_{\mathrm{D},\Sigma}(\boldsymbol{r}) = -\kappa^2 \int_\Sigma d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}') \boldsymbol{\nabla}_{r'} \phi_\Sigma(\boldsymbol{r}') + \int_\Gamma d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; \gamma_1^\Sigma \boldsymbol{F}_{\mathrm{D},\Sigma}(\boldsymbol{r}'), \quad \text{in } \Sigma\,,$$

and we finally obtain the reformulation of the correlation field $\boldsymbol{F}$ with Dirichlet boundary condition:

Theorem 4.3.3 *Reformulation of the correlation field $\boldsymbol{F}$ with Dirichlet boundary condition: The integral equation*

$$\boldsymbol{F}_{D,\Sigma}(\boldsymbol{r}) = -\kappa^2 \int_\Sigma d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}') \boldsymbol{\nabla}_{r'} \phi_\Sigma(\boldsymbol{r}') + \int_\Gamma d\Gamma_{r'}\, \gamma_{0,r'}\; \gamma_1^\Sigma \boldsymbol{F}_{D,\Sigma}(\boldsymbol{r}'), \quad \text{in } \Sigma$$

is the unique solution of the Dirichlet problem:

$$\begin{aligned} \mathcal{L}_\kappa \boldsymbol{F}_{D,\Sigma} &= \kappa^2 \boldsymbol{\nabla} \phi_\Sigma, && \text{in } \Sigma \\ \boldsymbol{F}_{D,\Sigma} &= \boldsymbol{0}, && \text{on } \Gamma \end{aligned}$$

Replacing the integral representation of the correlation field in the Lorentzian model, Theorem 3.3.1, by the Dirichlet system given in Theorem 4.3.3, we end up with Theorem 4.3.4.

Theorem 4.3.4 *The Lorentzian model of the biomolecular system in differential form with vanishing Dirichlet boundary condition:*

$$\begin{aligned} \varepsilon_\Omega \triangle \phi_\Omega &= -\rho, && \text{in } \Omega \\ \varepsilon_\infty \triangle \phi_\Sigma - (\varepsilon_\Sigma - \varepsilon_\infty)\,(\boldsymbol{\nabla} \cdot \boldsymbol{F}_{D,\Sigma}) &= 0, && \text{in } \Sigma \\ \phi_\Omega &= \phi_\Sigma && \text{on } \Gamma \\ \varepsilon_\Omega \partial_n \phi_\Omega &= \varepsilon_\infty \partial_n \phi_\Sigma - (\varepsilon_\Sigma - \varepsilon_\infty) \underbrace{(\boldsymbol{n} \cdot \boldsymbol{F}_{D,\Sigma})}_{=0}, && \text{on } \Gamma \\ \mathcal{L}_\kappa \boldsymbol{F}_{D,\Sigma} &= \kappa^2 \boldsymbol{\nabla} \phi_\Sigma, && \text{in } \Sigma \\ \boldsymbol{F}_{D,\Sigma} &= \boldsymbol{0}, && \text{on } \Gamma \end{aligned}$$

In the remainder of this work, this vector model for electrostatics in the biomolecular system is called the *Dirichlet Vector Model* (DVM).

Important remark 4.4 The incorporation of the boundary values into the definition of the correlation field $\boldsymbol{F}$ opens the possibility for a specific modulation of the dielectric effects on the boundary Γ. This extension, however, complicates the energy calculation. In the case of the Dirichlet approach we have to consider, for instance, the following variation:

$$\begin{aligned}
\delta W_{field} &= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \phi(\boldsymbol{r})\delta\rho(\boldsymbol{r}) = \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \boldsymbol{E}(\boldsymbol{r})\cdot\delta\boldsymbol{D}(\boldsymbol{r}) \\
&= \int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \boldsymbol{E}(\boldsymbol{r})\cdot\left(\int\limits_{\mathbb{R}^3} d\boldsymbol{r}'\, \varepsilon(\boldsymbol{r},\boldsymbol{r}')_{\mathbb{R}^3}\, \delta\boldsymbol{E}(\boldsymbol{r}') + \int\limits_{\Gamma} d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; \delta(\gamma_1^{\Sigma}\boldsymbol{F}_{\mathrm{D},\Sigma}(\boldsymbol{r}'))\right) \\
&= \underbrace{\frac{1}{2}\int\limits_{\mathbb{R}^3} d\boldsymbol{r}\delta(\boldsymbol{D}_{\mathrm{N}}(\boldsymbol{r})\cdot\boldsymbol{E}(\boldsymbol{r}))}_{(i)} + \underbrace{\int\limits_{\mathbb{R}^3} d\boldsymbol{r}\, \boldsymbol{E}(\boldsymbol{r})\cdot\left(\int\limits_{\Gamma} d\Gamma_{r'}\, \gamma_{0,r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\; \delta(\gamma_1^{\Sigma}\boldsymbol{F}_{\mathrm{D},\Sigma}(\boldsymbol{r}'))\right)}_{(ii)}
\end{aligned}$$

As we can see, the variation of the dielectric field $\delta\boldsymbol{D}$ cannot simply be shifted to the electric field $\boldsymbol{E}$ as it was done in Section 3.3.2.2, because we do not know *a priori* the variation of the boundary values $\gamma_1^{\Sigma}\boldsymbol{F}_{\mathrm{D},\Sigma}$, when increasing the charge $\delta\rho$. A final integration of the part (i) formally reveals the same contribution to the field energy as the Newton potential approach, see Theorem 3.3.2. The existence of the part (ii) is due to the incorporation of boundary effects - the maximization of the water correlations - into the Dirichlet approach. In general, the second contribution could account for specific polar and nonpolar regions that locally modify the water correlations. Having this in mind, the second term (ii) can be interpreted in parts as a cavity energy term that is commonly treated separately [8] (see Section 3.1.4.2).

In all the following calculations of the solvation free energy, we use the part (i) as the electrostatic contribution to the solvation free energy. The resulting energy W_{field} is obtained by a final integration as explained in Theorem 3.3.2. Calculating the field energy contribution of (ii) is one of the main aspects of future work.

4.3.2.1 The Born model in the Dirichlet approach

The spherical symmetry allows the gradient ansatz for all appearing fields, in particular

$$\boldsymbol{F}_{\mathrm{D},\Sigma} := -\boldsymbol{\nabla} F_{\mathrm{D},\Sigma}\,, \quad \text{in } \Sigma\,.$$

We solve the partial differential system given in Theorem 4.3.2 for a point charge, $\rho = q\,\delta(\boldsymbol{r})$ in the center of a sphere with radius a located at the origin.

$$\left.\begin{aligned}
\varepsilon_\Omega \triangle\phi_\Omega &= -q\,\delta(r) && r<a \\
\triangle[\varepsilon_\infty\phi_\Sigma + (\varepsilon_\Sigma-\varepsilon_\infty)F_{\mathrm{D},\Sigma}] &= 0 && r>a \\
\phi_\Omega &= \phi_\Sigma && r=a \\
\varepsilon_\Omega\partial_n\phi_\Omega &= \varepsilon_\infty\partial_n\phi_\Sigma && r=a \\
\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} &= -\kappa^2\phi_\Sigma && r>a \\
\partial_n F_{\mathrm{D},\Sigma} &= 0 && r=a
\end{aligned}\right\} \tag{4.24}$$

The analytical solution of system (4.24) is

$$\begin{aligned}
\phi_\Omega &= \frac{1}{4\pi}[\frac{1}{\varepsilon_\Omega\, r} + A] \\
\phi_\Sigma &= \frac{1}{4\pi r}\,[\frac{1}{\varepsilon_\Sigma} + B\, e^{-s\,(r-a)}] \\
F_\Sigma &= \frac{1}{4\pi r}\,[\frac{1}{\varepsilon_\Sigma} - B\, e^{-s\,(r-a)}\frac{\varepsilon_\infty}{\varepsilon_\Sigma-\varepsilon_\infty}]
\end{aligned}$$

where s, B and $B_2(B)$ are defined by

$$\begin{aligned} s &= \tfrac{1}{\lambda} = \sqrt{\tfrac{\varepsilon_\Sigma}{\varepsilon_\infty}}\kappa \\ A &= (\tfrac{1}{\varepsilon_\Sigma} - \tfrac{1}{\varepsilon_\Omega} + B)\tfrac{1}{a} \\ B &= (\tfrac{1}{\varepsilon_\infty} - \tfrac{1}{\varepsilon_\Sigma})\tfrac{1}{1+s\,a} \end{aligned}$$

4.4 The correlation field in differential form

As the reformulation of the integral representation of $\boldsymbol{F}$ is the key for all steps going further in application and theory, we now discuss some properties which are deduced from the differential form of $\boldsymbol{F}$:

Independent of the boundary values of $\boldsymbol{F}_\Sigma$, i.e., valid for $\boldsymbol{F}_{\mathrm{N},\Sigma}$ as well as for $\boldsymbol{F}_{\mathrm{D},\Sigma}$, the partial differential system to determine $\boldsymbol{F}_\Sigma$ and ϕ_Σ is given by:

$$\mathcal{L}_\kappa \boldsymbol{F}_\Sigma = \kappa^2 \boldsymbol{\nabla}\phi_\Sigma, \qquad \text{in } \Sigma \tag{4.25a}$$

$$\varepsilon_\infty \triangle\phi_\Sigma - (\varepsilon_\Sigma - \varepsilon_\infty)\,(\boldsymbol{\nabla}\cdot\boldsymbol{F}_\Sigma) = 0, \qquad \text{in } \Sigma \tag{4.25b}$$

Based on these equations we can extract general features of the nonlocal model in differential form:

Screening: the system (4.25) is characterized by a screening of the correlation field $\boldsymbol{F}_\Sigma$ on length scales $\sim 1/\kappa = \lambda\sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}$. Further, it is driven by the electrostatic field $-\boldsymbol{\nabla}\phi_\Sigma$, which plays the role of a source term. The appearance of $1/\kappa$ as characteristic length scale is surprising, since the Lorentzian model describes correlations on the scale λ. Only if the second equation is incorporated by the source term resulting in

$$\boldsymbol{\nabla}\cdot\left[\left(\triangle - \frac{1}{\lambda^2}\right)\boldsymbol{F}_\Sigma\right] = \boldsymbol{0}\,, \tag{4.26}$$

we obtain the physical length scale.

As we have seen, the screening in the case of spherical symmetry is only defined by λ. Although in general, both λ and $1/\kappa$ are length scales of the variations of the correlation field, this gives reason to believe that the contribution on length scales $1/\kappa$ is smaller in comparison to the one of the water network.

Limiting process: in Section 3.3.2, we have verified that the nonlocal Lorentzian model turns into a local response for $(\kappa \to \infty)$ and $(\kappa \to 0)$, respectively. This behavior is also reproduced in the differential formulation, i.e., *after* the application of the Yukawa operator:

$$\begin{aligned} \lim_{\kappa\to\infty} \quad & \left(\mathcal{L}_\kappa \boldsymbol{F}_\Sigma = \kappa^2 \boldsymbol{\nabla}\phi_\Sigma\right) \quad \Rightarrow \quad \boldsymbol{F}_\Sigma = -\boldsymbol{\nabla}\phi_\Sigma \quad \text{in } \Sigma \\ \lim_{\kappa\to 0} \quad & \left(\mathcal{L}_\kappa \boldsymbol{F}_\Sigma = \kappa^2 \boldsymbol{\nabla}\phi_\Sigma\right) \quad \Rightarrow \quad \boldsymbol{F}_\Sigma = \boldsymbol{0} \qquad\quad \text{in } \Sigma \end{aligned}$$

The local equations are revealed in Σ for both limiting processes as proved by the insertion of the limits in Eq. (4.25b).

In the case of a vanishing Dirichlet trace as proposed in Theorem 4.3.4, the transmission condition, however, remains unchanged

$$\varepsilon_\Omega \partial_n \phi_\Omega = \varepsilon_\infty \partial_n \phi_\Sigma \neq \varepsilon_\Sigma \partial_n \phi_\Sigma\,, \quad \text{on } \Gamma\,.$$

Although this conflicts with the local transmission condition, it is not a problem for our applications, since we always assume a finite correlation length.

4.5 Conclusion

In this chapter we deduced a general formulation of the nonlocal effect in the biomolecular system, which (a) is capable to incorporate a change of the water correlations in terms of boundary values of the correlation field $\boldsymbol{F}$ near the molecular surface Γ and (b) can be cast into a system of differential equations. This is necessary to easily solve the nonlocal equations for an arbitrary molecular geometry, because the electrostatic equations are also given in differential form.

The novel formulation holds new insight in the interpretation and the physical reasoning of the models given in [52] and [40], where the correlation field $\boldsymbol{F}$ is implicitly fixed on Γ: this is the Dirichlet Vector model with a correlation field $\boldsymbol{F}$ of vanishing Dirichlet trace on the molecular boundary Γ and the Newton Vector model, which is based on a globally continuous correlation field.

The proposed formulation goes qualitatively *beyond* these former models, because it offers the possibility to incorporate the specific nature of water on the molecular surface. In this way, the Dirichlet and the Newton Vector models are examples for a specially imposed boundary behavior.

The generality of the novel formulation has also been discussed in the light of energy contributions: a specification of the correlation field $\boldsymbol{F}$ on Γ gives rise to a cavity energy term, which originates from the reorganization of the water network when building the molecule's cavity. Although the additional energy contribution has not been further studied in this work, it clearly shows the capability of the presented formulation for nonlocal water models.

Chapter 5

Scalar approximations of the Lorentzian model

Although we eliminated the integral definition of the correlation field $\boldsymbol{F}$, the systems given in Theorem 4.3.2 and in Theorem 4.3.4 are still very complicated, because the unknowns are vectorial quantities. For a broad application of nonlocal electrostatics on biomolecules, the vectorial systems are impractical: the requirements on memory and computing time strongly limit the dimension of the biomolecules under consideration. Thus, the question is whether we can find simpler models, which capture the main features of nonlocal electrostatics.

A crucial simplification in the solving process is gained when the vectorial correlation field $\boldsymbol{F}$ can be expressed by the gradient of a potential field. In general, this is not the case, because the correlation field $\boldsymbol{F}$ as well as the dielectric permittivity $\boldsymbol{D}$ are not free of vortices

$$\boldsymbol{\nabla}\times\boldsymbol{D} \neq \boldsymbol{0} \qquad \text{and} \qquad \boldsymbol{\nabla}\times\boldsymbol{F} \neq \boldsymbol{0}\,, \quad \text{in } \mathbb{R}^3\,.$$

In general, $\boldsymbol{F}$ and $\boldsymbol{D}$ are therefore composed of a gradient and a rotational part, for instance, in the case of $\boldsymbol{F}$

$$\boldsymbol{F} = \boldsymbol{\nabla}F + \boldsymbol{\nabla}\times\boldsymbol{\xi}\,.$$

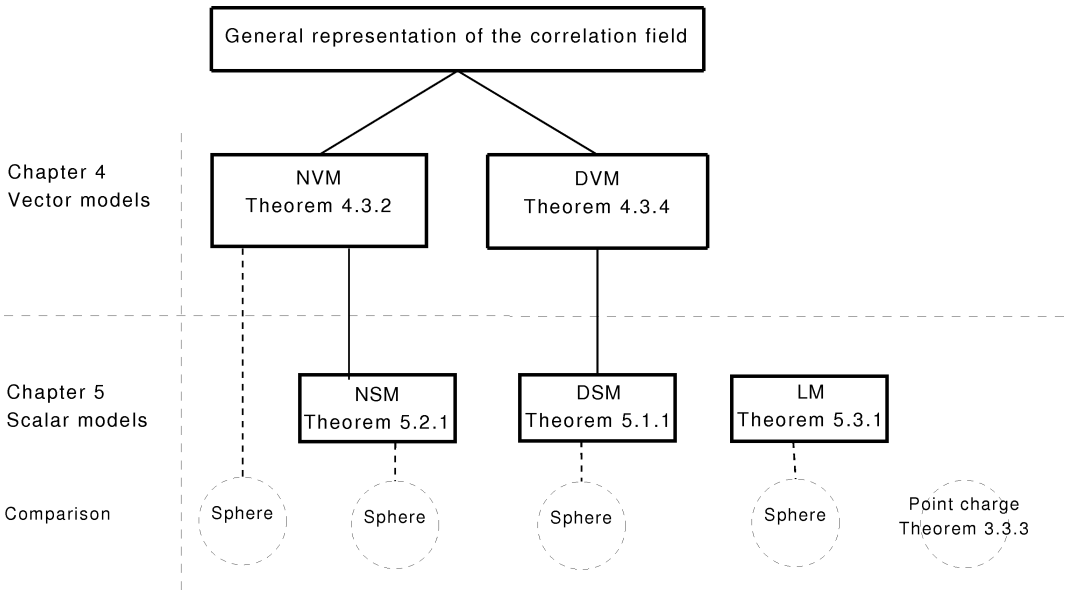


Fig. 5.1: Overview of two different scalar approaches for the Lorentzian water model in the biomolecular system deduced from the vector models.

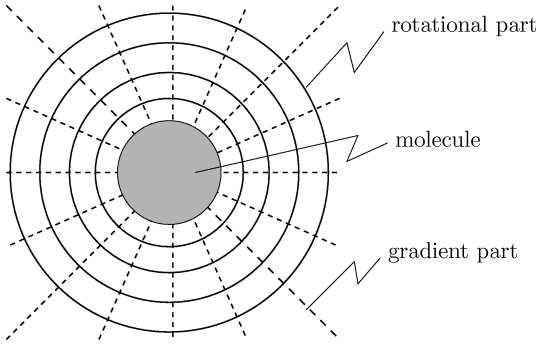


Fig. 5.2: Decomposition of a vector field $\boldsymbol{F} \in L^2(\mathbb{R}^3)$ into a rotational and a gradient part.

In the following sections, we pursue two different approaches to characterize the physical value of the rotational part $\boldsymbol{\nabla}\times\boldsymbol{\xi}$ of the correlation field $\boldsymbol{F}$.

The first approach starts with the Dirichlet Vector Model (DVM). We discuss the so called *Helmholtz decomposition* of the correlation field [29] (Section 5.1) and deduce a scalar model from the DVM, the so called Dirichlet Scalar Model (DSM).

The second ansatz approximates the Newton Vector Model (NVM) by gradient fields. Here, the potential ansatz,

$$\boldsymbol{F}_{\mathrm{N},\Sigma} = \boldsymbol{\nabla} F_{\mathrm{N},\Sigma}\,,$$

in fact, reduces to a permutation of the nonlocal dielectric function and the differential operator, which is *a priori* not given by first principles (Section 5.2). With the potential approach we argue for a second scalar model, the Newton Scalar Model (NSM).

An overview of the scalar models derived from the vector models is given in the diagram 5.1. A comparison on the Born sphere is given in Section 5.3. The models we take into account for this comparison are shown in the last row of Fig. 5.1.

5.1 Decomposition of the Dirichlet Vector Model

As part of the electrostatic material equations, it holds that the correlation field $\boldsymbol{F} \in L^2(\mathbb{R}^3)^3$. This allows to consider $\boldsymbol{F}$ in a Helmholtz decomposition (see [29] p. 314). The Helmholtz decomposition describes the mathematical concept of decomposing a vector field $\boldsymbol{F} \in L^2(V)^3$ in a domain V in two orthogonal contributions:

$$\begin{aligned}
&\boldsymbol{F} = \boldsymbol{\nabla} F \oplus \boldsymbol{\nabla}\times\boldsymbol{\xi}\,, \quad \text{in } V \\
&\boldsymbol{\nabla}\times(\boldsymbol{\nabla} F) = \boldsymbol{0} && \rightarrow \text{ gradient part is curl-free} \\
&\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\boldsymbol{\xi}) = 0 && \rightarrow \text{ rotational part is divergence-free}
\end{aligned}$$

Fig. 5.2 exemplifies a curl-free and a divergence-free vector field. The dashed lines correspond to the streamlines of the curl-free and the circles to the streamlines of the divergence-free vector field. The Helmholtz decomposition is a good starting point for the analysis of the rotational and the gradient part of the correlation field $\boldsymbol{F}$.

Formally, the Helmholtz decomposition of $L^2(V)^3$ in an arbitrary bounded [1], simply connected

[1] An analogous decomposition holds for unbounded regions with bounded complements, e.g., for Σ in the biomolecular system. The Sobolev space H^1 then has to be substituted by the so called Beppo-Levi space W^1 [29].

domain $V \in \mathbb{R}^3$ can be written as follows

$$L^2(V)^3 = \overbrace{H_0(\mathrm{curl}\,0,V)}^{\text{gradient part}} \oplus \overbrace{\boldsymbol{\nabla}\times\ H^1(V)^3}^{\text{rotational part}} \tag{5.1a}$$

$$L^2(V)^3 = \underbrace{H_0(\mathrm{curl}\,0,V) \oplus \mathrm{L}^2_{rg}(V)^3}\ \oplus H_0(\mathrm{div}\,0,V) \tag{5.1b}$$

$$L^2(V)^3 = \underbrace{\boldsymbol{\nabla} H^1(V)}_{\text{gradient part}} \oplus \underbrace{H_0(\mathrm{div}\,0,V)}_{\text{rotational part}}, \tag{5.1c}$$

where the definition of the different spaces is given in Appendix 10.1.3. The operator $\oplus$ implies the orthogonality of the decomposition [29]. The attribute "orthogonality" here means that a function of one space can never be represented by a linear combination of functions in the remaining orthogonal spaces or in other words, two functions belonging to different orthogonal spaces cannot cancel each other.

The three decompositions split the space $L^2(V)^3$ into slightly different ways: the most detailed, orthogonal decomposition is given in Eq. (5.1b), where the decision to represent the contribution of $\mathrm{L}^2_{rg}(V)^3$ as gradient of a scalar potential like in Eq. (5.1a) or as the curl of a vectorial potential like in Eq. (5.1c) is not made. This is shown by the two central braces which clarify the connectivity of $\mathrm{L}^2_{rg}(V)^3$ to be part of $\boldsymbol{\nabla} H^1(V)$ or $\boldsymbol{\nabla}\times\ H^1(V)^3$.

Our aim is to find a pure gradient approach, and therefore we want to justify models with negligible rotational part. This, of course, implies the reduction of the solution space. Thus, it is reasonable to analyze an orthogonal decomposition, where the rotational part captures the smallest possible part of $L^2(V)^3$. This means that we have to start with decomposition (5.1c) for the following discussion.

In the decomposition

$$L^2(V)^3 = \underbrace{\boldsymbol{\nabla} H^1(V)}_{\text{gradient part}} \oplus \underbrace{H_0(\mathrm{div}\,0,V)}_{\text{rotational part}},$$

the rotational part has a fixed behavior at the boundary ∂V of V, because $H_0(\mathrm{div}\,0,V)$ is defined as follows

$$H_0(\mathrm{div}\,0,V) = \{\boldsymbol{v} \in L^2(V)^3, \boldsymbol{\nabla}\cdot\boldsymbol{v} = 0 \text{ in } V,\ \boldsymbol{n}\cdot\boldsymbol{v} = 0 \text{ on } \partial V\}.$$

This means that the normal component vanishes like it is the case for the rotational vector field in Fig. 5.2. In contrast, the boundary values of the gradient field remain undetermined.

We define the orthogonal decomposition [2] of the correlation field $\boldsymbol{F}_{\mathrm{D},\Sigma}$ of the Dirichlet Vector Model in Theorem 4.3.4:

$$L^2(\Sigma)^3 \ni \boldsymbol{F}_{\mathrm{D},\Sigma} = -\boldsymbol{\nabla} F_{\mathrm{D},\Sigma} + \boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma, \quad \text{in } \Sigma \tag{5.2}$$
$$\text{with} \qquad -\boldsymbol{\nabla} F_{\mathrm{D},\Sigma} \in \boldsymbol{\nabla} H^1(\Sigma)$$
$$\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma \in H_0(\mathrm{div}\,0,\Sigma),$$

where $\boldsymbol{\nabla} F_{\mathrm{D},\Sigma}$ and $\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma$ decrease fast enough in the limit $|\boldsymbol{r}| \to \infty$ so that the outer boundary integrals vanish [29], i.e., the vector fields fulfill the radiation condition. Replacing $\boldsymbol{F}_{\mathrm{D},\Sigma}$ with the

[2] $\int_\Sigma dV\,(\boldsymbol{\nabla} F_{\mathrm{D},\Sigma})\cdot(\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma) = -\int_\Sigma dV\ \underbrace{(\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma))}_{=0}\ F_{\mathrm{D},\Sigma} + \int_\Gamma d\Gamma\ \underbrace{[\boldsymbol{n}\cdot(\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma)]}_{=0}\ F_{\mathrm{D},\Sigma} = 0,$

ansatz (5.2) in Theorem 4.3.4 yields

$$\begin{aligned}
\varepsilon_\Omega \triangle \phi_\Omega &= -\rho, && \text{in } \Omega && \text{(5.3a)}\\
\triangle[\varepsilon_\infty \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty) F_{\mathrm{D},\Sigma}] &= 0, && \text{in } \Sigma && \text{(5.3b)}\\
\varepsilon_\Omega \partial_n \phi_\Omega &= \varepsilon_\infty \partial_n \phi_\Sigma, && \text{on } \Gamma && \text{(5.3c)}\\
\phi_\Omega &= \phi_\Sigma, && \text{on } \Gamma && \text{(5.3d)}\\
\mathcal{L}_\kappa(-\boldsymbol{\nabla} F_{\mathrm{D},\Sigma} + \boldsymbol{\nabla}\times \boldsymbol{\xi}_\Sigma) &= \kappa^2 \boldsymbol{\nabla}\phi_\Sigma, && \text{in } \Sigma && \text{(5.3e)}\\
\partial_n F_{\mathrm{D},\Sigma} &= 0, && \text{on } \Gamma && \text{(5.3f)}\\
\boldsymbol{n}\times(-\boldsymbol{\nabla} F_{\mathrm{D},\Sigma} + \boldsymbol{\nabla}\times \boldsymbol{\xi}_\Sigma) &= \mathbf{0}, && \text{on } \Gamma\,, && \text{(5.3g)}
\end{aligned}$$

where we explicitly used that $\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma \in H_0(\operatorname{div} 0, \Sigma)$ and thus

$$\boldsymbol{n}\cdot(\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma) = 0\,, \quad \text{on } \Gamma\,. \tag{5.4}$$

A closer look at system (5.3) shows that the choice of the decomposition allowed us to separate the rotational from the gradient part in the Neumann boundary condition of the field $F_{\mathrm{D},\Sigma}$, Eq. (5.3f), and in the Neumann transmission condition of the electric field, Eq. (5.3c). Further, we see that the material equations are determined by the gradient part of the correlation field only. However, the rotational part still appears in Eq. (5.3e) and Eq. (5.3g). Eq. (5.3e) might be further simplified, if we can prove the orthogonality of the fields $\big(\boldsymbol{\nabla}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma)\big)$ and $(\boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma))$:

$$\begin{aligned}
&\int_\Sigma dV\ \big(\boldsymbol{\nabla}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma)\big)\cdot(\boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma))\\
&\overset{\text{Eq. (5.3e)}}{=} \int_\Sigma dV\ (\boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma))\cdot(\boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma)) = \int_\Sigma dV\ |\boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma)|^2\\
&\overset{!}{=} \quad 0 \quad \Leftrightarrow \quad \boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma) = \mathbf{0}
\end{aligned} \tag{5.5}$$

Eq. (5.5) tells us that in case of orthogonality, Eq. (5.3e) separates into the following equations,

$$\begin{aligned}
\boldsymbol{\nabla}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma) &= \mathbf{0}\\
\boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma) &= \mathbf{0}\,.
\end{aligned}$$

Unfortunately, the orthogonality is not guaranteed in general as shown by the following calculation:

$$\begin{aligned}
&\int_\Sigma dV\ \big(\boldsymbol{\nabla}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma)\big)\cdot(\boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma))\\
&\quad= -\int_\Sigma dV\ \underbrace{[\boldsymbol{\nabla}\cdot(\boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma))r]}_{=0}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma)\\
&\qquad+\int_\Gamma d\Gamma\,[\boldsymbol{n}\cdot(\boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma))]\,(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma)\\
&\quad= \frac{1}{\varepsilon_\infty}\int_\Gamma d\Gamma\ \underbrace{[\boldsymbol{n}\cdot(\boldsymbol{\nabla}\times(\triangle\boldsymbol{\xi}_\Sigma))]}_{\text{i.g. } \neq 0}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma)\,,
\end{aligned} \tag{5.6}$$

In the next two paragraphs we motivate to search for solutions where the orthogonality is conserved.

5.1.1 Screening length of orthogonal solutions

First, we have a look at the differential equations of system (5.3) in Σ.

$$\begin{aligned} \triangle[\varepsilon_\infty\phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty)F_{\mathrm{D},\Sigma}] &= 0, && \text{in } \Sigma \\ \boldsymbol{\nabla}\left(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma\right) &= \boldsymbol{\nabla}\times(\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma), && \text{in } \Sigma \end{aligned}$$

If we, for the moment, assume the rotational part $\mathcal{L}_\kappa(\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma)$ to vanish separately, we can apply the divergence operator $(\boldsymbol{\nabla}\cdot)$ on the latter equation without the loss of any degree of freedom and obtain

$$\left\{ \begin{aligned} \triangle[\varepsilon_\infty\phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty)F_{\mathrm{D},\Sigma}] &= 0, && \text{in } \Sigma \\ \mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma &= 0, && \text{in } \Sigma \end{aligned} \right\} \tag{5.7}$$

$$\Leftrightarrow \qquad \left\{ \triangle\left[(\triangle - \frac{1}{\lambda^2})F_{\mathrm{D},\Sigma}\right] = 0, \quad \text{in } \Sigma \right\}. \tag{5.8}$$

We see from system (5.7) that ϕ_Σ is driven by $F_{\mathrm{D},\Sigma}$. This source term is characterized by the length scale λ as apparent in Eq. (5.8). It is due to an incorporation of nonlocal correlations that, by definition, decrease on the scale λ.

In contrast, if we assume $\mathcal{L}_\kappa(\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma)$ not to be orthogonal to $\boldsymbol{\nabla}\left(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma\right)$, the physically reasonable assumption of the screening length λ is no more guaranteed; a screening on the length scale $1/\kappa$ is allowed as well. This means that the rotational part induces variations of the electrostatic potential on a length scale which we assume is of minor importance in comparison with the one of the water network. In the case of spherical symmetry, this contribution actually vanishes, see Remark 4.3.

5.1.2 Existence of orthogonal solutions

From Eq. (5.6) we know that the orthogonality of $\boldsymbol{\nabla}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma)$ and $(\boldsymbol{\nabla}\times\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma)$ is given for

$$\boldsymbol{n}\cdot(\mathcal{L}_\kappa\boldsymbol{\nabla}\times\boldsymbol{\xi}_\Sigma) = 0\,, \quad \text{on } \Gamma\,.$$

In the following, we analyze this requirement in the special case of a surface, which can locally be assumed to be part of a sphere as it is the case for a biomolecule almost everywhere. Exploiting this fact, we describe the vector field $\boldsymbol{\xi}$ in a small region around $\boldsymbol{r}\in\Gamma$ in the spherical coordinate system that is located in the origin of the sphere whose surface locally coincides with the surface of the molecule. Then, the surface normal direction is given by the radial unit vector of the spherical coordinate system and the complementing orthogonal unit vectors correspond to the tangents on the molecule's surface. This is illustrated in Fig. 5.3 for two different situations.

With these assumptions, the vector field $\boldsymbol{\xi}$ can be locally written as

$$\boldsymbol{\xi}(\boldsymbol{r}) = \xi_r(r,\theta,\phi)\,\boldsymbol{e}_r + \xi_\theta(r,\theta,\phi)\,\boldsymbol{e}_\theta + \xi_\phi(r,\theta,\phi)\,\boldsymbol{e}_\phi\,, \quad \text{in } \Sigma\,,$$

where we know from the chosen Helmholtz decomposition that

$$\boldsymbol{n}\cdot[(\boldsymbol{\nabla}\times\boldsymbol{\xi})] = \boldsymbol{e}_r\cdot[(\boldsymbol{\nabla}\times\boldsymbol{\xi})] = \frac{1}{r\sin\theta}(\partial_\theta(\sin\theta\,\xi_\phi(r,\theta,\phi)) - \partial_\phi\xi_\theta(r,\theta,\phi)) = 0\,, \quad \text{on } \Gamma. \tag{5.9}$$

The application of the operator $(\triangle\,(\boldsymbol{\nabla}\times))$ and the projection on $\boldsymbol{n}$ reveals

$$\boldsymbol{n}\cdot[\triangle\,(\boldsymbol{\nabla}\times\boldsymbol{\xi})] = \boldsymbol{e}_r\cdot[\triangle\,(\boldsymbol{\nabla}\times\boldsymbol{\xi})] = \boldsymbol{e}_r\cdot[\triangle\,(\underbrace{(\boldsymbol{\nabla}\times\boldsymbol{\xi})_r\boldsymbol{e}_r}_{(*)} + (\boldsymbol{\nabla}\times\boldsymbol{\xi})_\theta\boldsymbol{e}_\theta + (\boldsymbol{\nabla}\times\boldsymbol{\xi})_\phi\boldsymbol{e}_\phi)]\,, \quad \text{on } \Gamma\,. \tag{5.10}$$

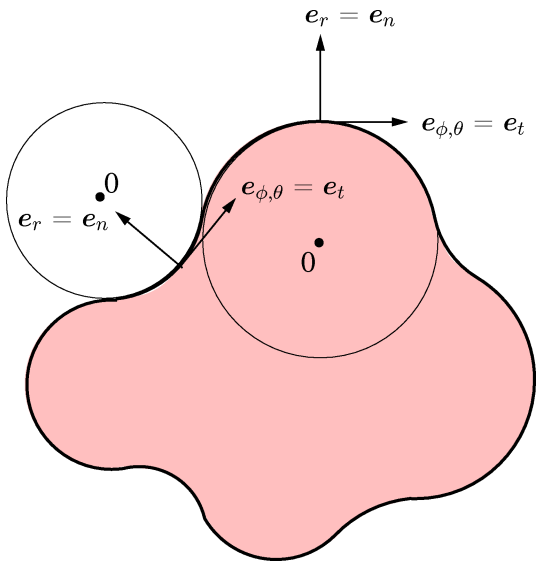


Fig. 5.3: The biomolecular surface is locally well described by spherical coordinates.

It is physically reasonable to assume that $(*)$ is zero not only for the limiting process, but also in a small domain around Γ, as we argued that the correlation field

$$\boldsymbol{F}_{\mathrm{D},\Sigma} = \boldsymbol{\nabla} F_{\mathrm{D},\Sigma} + \boldsymbol{\nabla}\times\boldsymbol{\xi}\,,$$

vanishes on Γ and smoothly increases in magnitude (favorably in normal direction). In particular, this holds for the rotational part with its vanishing normal component on Γ. Thus, we set the expression $(*)$ to zero before the limiting process, $\Sigma \ni \boldsymbol{r} \to \Gamma$, and further simplify Eq. (5.10):

$$\boldsymbol{n}\cdot[\triangle(\boldsymbol{\nabla}\times\boldsymbol{\xi})] = \boldsymbol{e}_r\cdot[\triangle(\boldsymbol{\nabla}\times\boldsymbol{\xi})] = \frac{2}{r^3\sin\theta}(r\partial_r\left[\partial_\theta(\sin\theta\,\xi_\phi(r,\theta,\phi)) - \partial_\phi\xi_\theta(r,\theta,\phi)\right]),\ \text{on}\,\Gamma\,, \quad (5.11)$$

where we used the following identities

$$\begin{aligned}(\boldsymbol{\nabla}\times\boldsymbol{\xi})_\theta &= \frac{1}{r\,\sin\theta}\left(\partial_\phi\xi_r(r,\theta,\phi) - \partial_r(r\sin\theta\,\xi_\phi(r,\theta,\phi))\right)\\ (\boldsymbol{\nabla}\times\boldsymbol{\xi})_\phi &= \frac{1}{r}\left(\partial_r(r\xi_\theta(r,\theta,\phi)) - \partial_\theta\xi_r(r,\theta,\phi)\right)\,.\end{aligned}$$

Now, we can estimate the approximations required for preserving the orthogonality of the two fields, $\boldsymbol{\nabla}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2\phi_\Sigma)$ and $(\boldsymbol{\nabla}\times\mathcal{L}_\kappa\boldsymbol{\xi}_\Sigma)$: this is the case, when the expression (5.11) vanishes.

Together with Eq. (5.9) we see that this is fulfilled for the separation ansatz

$$\begin{aligned}\xi_\theta(r,\theta,\phi) &= b(r)\,\xi_\theta(\theta,\phi)\\ \xi_\phi(r,\theta,\phi) &= b(r)\,\xi_\phi(\theta,\phi)\,.\end{aligned}$$

This ansatz implies that the dependence on r is equal for ξ_θ and ξ_ϕ and further that it is locally independent from variations in θ and ϕ. In a first approximation, such an assumption is reasonable as both tangential parts have to decrease with increasing distance to the molecular surface independent from the particular choice of θ and ϕ. In fact, this is the approximation we accept in order to obtain a gradient formulation.

5.1.3 Dirichlet Scalar Model

The discussion of the last sections reveals that we can find solutions where the orthogonality of the vector fields is conserved after the application of the Yukawa operator. These orthogonal solutions preserve the screening length λ as the only length scale of nonlocal correlations. Therefore, we now consider solutions of Eq. (5.3e) in Theorem 4.3.4, which split into two parts:

$$\begin{aligned} \boldsymbol{\nabla}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2 \phi_\Sigma) &= \mathbf{0}, \quad \text{in } \Sigma \quad \Rightarrow \qquad \mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2 \phi_\Sigma = 0, \quad \text{in } \Sigma \\ \boldsymbol{\nabla}\times \mathcal{L}_\kappa \boldsymbol{\xi}_\Sigma &= \mathbf{0}, \quad \text{in } \Sigma \end{aligned}$$

Incorporating this restriction in system (5.3) results in

$$\begin{aligned} \varepsilon_\Omega \triangle \phi_\Omega &= -\rho, && \text{in } \Omega \\ \triangle[\varepsilon_\infty \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty) F_{\mathrm{D},\Sigma}] &= 0, && \text{in } \Sigma \\ \phi_\Omega &= \phi_\Sigma, && \text{on } \Gamma \\ \varepsilon_\Omega \partial_n \phi_\Omega &= \varepsilon_\infty \partial_n \phi_\Sigma, && \text{on } \Gamma \\ \\ \mathcal{L}_\kappa F_{\mathrm{D},\Sigma} &= -\kappa^2 \phi_\Sigma, && \text{in } \Sigma \\ \partial_n F_{\mathrm{D},\Sigma} &= 0, && \text{on } \Gamma \\ [\, -\boldsymbol{n}\times(\boldsymbol{\nabla} F_{\mathrm{D},\Sigma}) &= \text{jump}, && \text{on } \Gamma\,] \,. \end{aligned}$$

Although we justified the orthogonality of the following fields

$$\boldsymbol{\nabla}(\mathcal{L}_\kappa F_{\mathrm{D},\Sigma} + \kappa^2 \phi_\Sigma) \oplus (\boldsymbol{\nabla}\times \mathcal{L}_\kappa \boldsymbol{\xi}_\Sigma)\,, \quad \text{in } \Sigma\,,$$

the rotational part still influences the tangential transmission condition on Γ. However, the scalar system is uniquely defined by the vanishing Neumann boundary condition. The jump in the tangential derivatives of $F_{\mathrm{D},\Sigma}$ as indicated in the last equation is implicitly determined by the differential system. We end up in the Dirichlet Scalar Model.

Theorem 5.1.1 *Dirichlet Scalar Model (DSM)*

$$\begin{aligned} \varepsilon_\Omega \Delta \phi_\Omega &= -\rho, && \text{in } \Omega \\ \triangle(\varepsilon_\infty\, \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty)\, F_{D,\Sigma}) &= 0, && \text{in } \Sigma \\ \phi_\Omega &= \phi_\Sigma, && \text{in } \Gamma \\ \varepsilon_\Omega \partial_n \phi_\Omega &= \varepsilon_\infty \partial_n \phi_\Sigma, && \text{in } \Gamma \\ \\ \mathcal{L}_\kappa F_{D,\Sigma} &= -\kappa^2 \phi_\Sigma, && \text{in } \Sigma \\ \partial_n F_{D,\Sigma} &= 0, && \text{in } \Gamma \end{aligned}$$

A comparison to the Dirichlet Vector Model (DVM) in the spherically symmetric case shows that it coincides with the Dirichlet Scalar Model (DSM). However, the derivation and argumentation for the DSM, which we gave in the previous sections, was more general and supports that this model yields physically reasonable results for non-trivial geometries as well.

5.2 Decomposition of the Newton Vector Model

In this section, we focus on the Newton Vector Model and try to reduce the vectorial equations to a set of scalar quantities. To this end, we compare the definition of $\boldsymbol{F}_\mathrm{N}$ in Eq. (4.22),

$$\begin{aligned} -\boldsymbol{F}_\mathrm{N}(\boldsymbol{r}) = &\quad \kappa^2 \int_\Sigma d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}') \quad \boldsymbol{\nabla}_{r'}\phi(\boldsymbol{r}')\,, \quad \text{in } \mathbb{R}^3 \\ \text{with} \qquad \boldsymbol{\nabla} f_\mathrm{N}(\boldsymbol{r}) := &\quad \boldsymbol{\nabla}\kappa^2 \int_\Sigma d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}') \quad \phi(\boldsymbol{r}')\,, \quad \text{in } \mathbb{R}^3\,. \end{aligned} \tag{5.12}$$

The similarity motivates to analyse the difference in the corresponding differential equations, i.e., when applying the Yukawa operator $\mathcal{L}_\kappa$. In Σ, the application reveals, for instance:

$$\begin{aligned} \boldsymbol{\nabla}\mathcal{L}_\kappa f_{\mathrm{N},\Sigma} &= -\kappa^2\,\boldsymbol{\nabla}\phi_\Sigma(\boldsymbol{r}), \quad \text{in } \Sigma \\ \mathcal{L}_\kappa \boldsymbol{F}_{\mathrm{N},\Sigma} &= -\kappa^2\,\boldsymbol{\nabla}\phi_\Sigma(\boldsymbol{r}), \quad \text{in } \Sigma \end{aligned}$$

Comparing these equations, we find that

$$\mathcal{L}_\kappa(\boldsymbol{F}_{\mathrm{N},\Sigma} - \boldsymbol{\nabla} f_{\mathrm{N},\Sigma}) = 0, \quad \text{in } \Sigma\,.$$

With the same argumentation, we can further prove that an analogous equation holds in Ω. This observation leads to the following result: the vector field $\boldsymbol{F}_\mathrm{N}$ can be decomposed into a gradient field, $-\boldsymbol{\nabla} f_\mathrm{N}$, and a function which lies in the kernel of the Yukawa operator $\mathcal{L}_\kappa$,

$$\Rightarrow \qquad \boldsymbol{F}_\mathrm{N} = -\boldsymbol{\nabla} f_\mathrm{N} + \boldsymbol{F}_\mathrm{N}^{hom} \qquad \text{with} \qquad \mathcal{L}_\kappa \boldsymbol{F}_\mathrm{N}^{hom} = 0\,, \quad \text{in } \Omega \cup \Sigma\,. \tag{5.13}$$

Remark 5.1 The gradient field $\boldsymbol{\nabla} f_{\mathrm{N},\Sigma}$ should not be confused with the gradient field $\boldsymbol{\nabla} F_{\mathrm{D},\Sigma}$. The latter is an element of the space $\boldsymbol{\nabla} H^1(\Sigma)$ and with its rotational counterpart $\boldsymbol{\nabla}\times\xi \in H_0(\mathrm{div}\,0,\Sigma)$ it completes the Helmholtz decomposition of a vectorial function in $L^2(\Sigma)^3$. A Helmholtz decomposition is not carried out for $\boldsymbol{F}_\mathrm{N}$ meaning that we did not decide whether the part of the solution in $L^2_{rg}(\Sigma)$ contributes to the field $\boldsymbol{\nabla} f_{\mathrm{N},\Sigma}$ or $\boldsymbol{F}_\mathrm{N}^{hom}$.

5.2.1 Characterization of $\mathbf{F}_\mathbf{N}^{hom}$

In order to characterize the physical meaning of $\boldsymbol{F}_\mathrm{N}^{hom}$ and therefore to estimate its impact on the nonlocal theory, we now try to find the reason for its appearance. To this end, we start with the (global) definition of the correlation field $\boldsymbol{F}_\mathrm{N}$ (Eq. (4.22)) and try to permute the differential and the integral operator:

$$\begin{aligned} \boldsymbol{F}_\mathrm{N}(\boldsymbol{r}) &= -\kappa^2 \int_{\mathbb{R}^3} d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\,\chi_\Sigma \boldsymbol{\nabla}_{r'}\phi_\Sigma(\boldsymbol{r}') = -\kappa^2 \int_\Sigma d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\,\boldsymbol{\nabla}_{r'}\phi_\Sigma(\boldsymbol{r}') \\ &= \kappa^2 \int_\Sigma d\boldsymbol{r}' \left(\boldsymbol{\nabla}_{r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\right)\phi_\Sigma(\boldsymbol{r}') - \kappa^2 \int_\Sigma d\boldsymbol{r}' \boldsymbol{\nabla}_{r'}\left(G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\phi_\Sigma(\boldsymbol{r}')\right) \\ &= \underbrace{-\kappa^2 \boldsymbol{\nabla} \int_\Sigma d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\phi_\Sigma(\boldsymbol{r}')}_{\boldsymbol{\nabla} f_\mathrm{N}} + \underbrace{\kappa^2 \int_\Gamma d\Gamma_{r'}\,\boldsymbol{n}\left(G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\phi_\Sigma(\boldsymbol{r}')\right)}_{\boldsymbol{F}_\mathrm{N}^{hom}}, \text{ in } \mathbb{R}^3\,, \end{aligned} \tag{5.14}$$

where the outer boundary integral has been directly set to zero because of the radiation condition. In Eq. (5.14) we identify $\boldsymbol{F}_{\mathrm{N}}^{hom}$ with a surface integral, which indeed has no contribution to the differential equations, because $\boldsymbol{r} \in \Omega \cup \Sigma$ and $\boldsymbol{r}' \in \Gamma$. Thus, the δ-distribution yields zero in the integral expression:

$$\begin{aligned} \mathcal{L}_\kappa \boldsymbol{F}_{\mathrm{N}}(\boldsymbol{r}) &= -\boldsymbol{\nabla}\mathcal{L}_\kappa f_{\mathrm{N}}(\boldsymbol{r}) - \kappa^2 \int_\Gamma d\Gamma_{r'}\, \mathcal{L}_\kappa\big(G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\big)\,\phi_\Sigma(\boldsymbol{r}') \\ &= -\boldsymbol{\nabla}\mathcal{L}_\kappa f_{\mathrm{N}}(\boldsymbol{r}) + \underbrace{\kappa^2 \int_\Gamma d\Gamma_{r'} \delta(\boldsymbol{r}-\boldsymbol{r}')\phi_\Sigma(\boldsymbol{r}')}_{=0}\,, \qquad \boldsymbol{r} \in \Omega \cup \Sigma \end{aligned}$$

In Section 4.3, we learned that surface integrals in the representation formula incorporate discontinuities of the corresponding field on the surface. Having this in mind, the difference of $\boldsymbol{F}_{\mathrm{N}}$ and $\boldsymbol{\nabla} f_{\mathrm{N}}$ originates from a different behavior on Γ. The vectorial Newton potential $\boldsymbol{F}_{\mathrm{N}}$ is defined by

$$\boldsymbol{\rho}(\boldsymbol{r}) = -\kappa^2 \chi_\Sigma \boldsymbol{\nabla}\phi_\Sigma(\boldsymbol{r}) = \begin{cases} \boldsymbol{0}\,, & \text{in } \Omega \\ -\kappa^2 \boldsymbol{\nabla}\phi_\Sigma(\boldsymbol{r})\,, & \text{in } \Sigma\,, \end{cases} \tag{5.15}$$

whereas the scalar Newton potential f_{N} has the following discontinuous integrand

$$\rho(\boldsymbol{r}) = -\kappa^2 \chi_\Sigma \phi_\Sigma(\boldsymbol{r}) = \begin{cases} 0\,, & \text{in } \Omega \\ -\kappa^2 \phi_\Sigma(\boldsymbol{r})\,, & \text{in } \Sigma\,. \end{cases}$$

Since the gradient operator $\boldsymbol{\nabla}$ does not commute with the characteristic function χ_Σ, the permutation of them yields a surface integral. The existence of the characteristic function in Eq. (5.15) has been interpreted as the incorporation of the sudden jump into the nonlocal dielectric function when crossing Γ. However, when we think of the flexibility of the polar side chains lying on the surface of the molecule, and when we think of a possible decrease in the hydrogen bonds due to the surface, the sudden jump from the local to the nonlocal dielectric function is a simplification and it is legitimate to consider whether a *smooth* extension of $-\kappa^2 \boldsymbol{\nabla}\phi_\Sigma$ to $\mathbb{R}^3$ would be a better choice. In the following section, we introduce the convolution as a means to smooth discontinuous functions. This concept is then applied in Section 5.2.3 on the Newton Vector Model for nonlocal electrostatics.

5.2.2 Convolutions as "smoothing" operations

Due to general features of convolutions [28], convoluting a discontinuous function with a so-called *mollifier* function is a mean to smooth the original discontinuous function [28, 55]. In order to illustrate this, we now convolute the discontinuous, truncated $\frac{1}{r}$-potential,

$$\forall\, \boldsymbol{r} \in \mathbb{R}^3: \quad \phi_{\Sigma,Born}(\boldsymbol{r}) = \begin{cases} 0 & |\boldsymbol{r}| = r < 1 \\ \frac{1}{1-r} & |\boldsymbol{r}| = r > 1 \end{cases}$$

with the mollifier function $s_e(\boldsymbol{r})$:

$$\forall\, \boldsymbol{r} \in \mathbb{R}^3: \quad s_e(\boldsymbol{r}) = \begin{cases} c_e\, e^{-1/(1-(\frac{r}{e})^2)} & |\boldsymbol{r}| = r < e \\ 0 & |\boldsymbol{r}| = r \geq e \end{cases}$$

where c_e is a normalization constant. The convolution $(s_e * \phi)$ can be described as a weighted

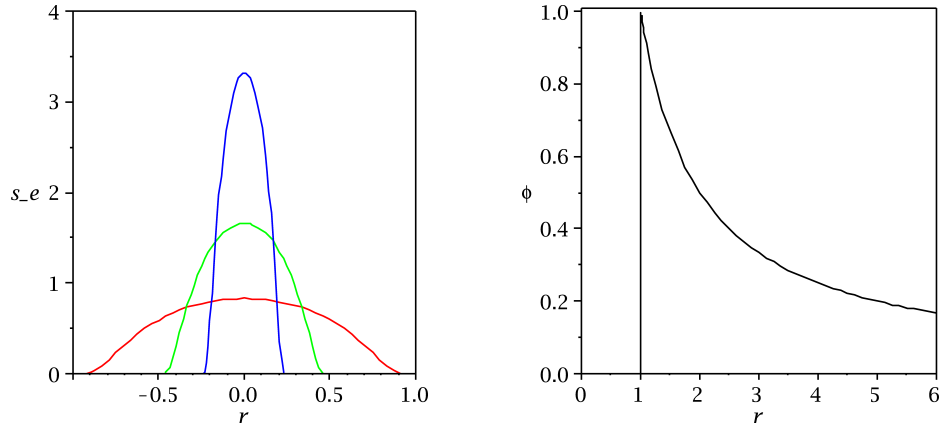


Fig. 5.4: (left) different mollifiers s_e ($e = 1$ =red; $e = 1/2$ =green; $e = 1/4$ =blue); (right) a discontinuous function $\phi_{\Sigma,Born}$ with compact support, which should be smoothed in the jump region.

average of the function ϕ with characteristic weighting given by the mollifier s_e. The parameter e coordinates the extent of the smoothing region as shown in Fig. 5.4. In the limit $e \to 0$, we end up in the δ-distribution, which - convoluted with the discontinuous function $\phi_{\Sigma,Born}$ - returns the original function $\phi_{\Sigma,Born}$. Fig. 5.5 shows the results of the following convolutions

$$(s_e * \phi)(\boldsymbol{r}) = \int_{\mathbb{R}^3} d\boldsymbol{r}' s_e(\boldsymbol{r} - \boldsymbol{r}')\, \phi(\boldsymbol{r}')\,, \quad \boldsymbol{r} \in \mathbb{R}^3\,; \quad \text{with } e \in \{1, \tfrac{1}{2}, \tfrac{1}{4}\}\,.$$

Compared to the discontinuous function $\phi_{\Sigma,Born}$, the convolutions exhibit a smooth transition in the jump region – of course at the expense of accuracy, assuming that a discontinuous jump was indeed physical. However, the smaller e becomes the better the agreement.

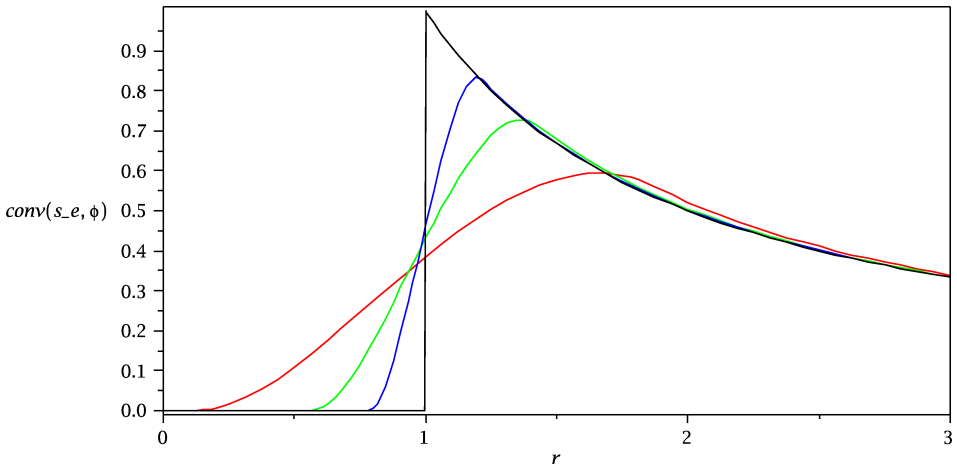


Fig. 5.5: Convolution of the mollifier s_e with $\phi_{\Sigma,Born}$ (black); ($e = 1$ =red; $e = 1/2$ =green; $e = 1/4$ = blue).

5.2.3 Newton Scalar Model

The study of the previous section motivates a modification of the discontinuous kernel of the correlation field $\boldsymbol{F}_{\mathrm{N}}$,

$$\boldsymbol{F}_{\mathrm{N}}(\boldsymbol{r}) = -\kappa^2 \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\,\boldsymbol{\rho}(\boldsymbol{r}'), \quad \text{in } \mathbb{R}^3$$

to a purely smooth potential approach. This means that we want to consider a smoothing operation on the truncated potential ϕ: instead of Eq. (4.21),

$$\boldsymbol{\rho}(\boldsymbol{r}) = -\kappa^2 \chi_\Sigma\, \boldsymbol{\nabla}\phi_\Sigma(\boldsymbol{r}), \quad \text{in } \mathbb{R}^3\,,$$

we consider

$$\tilde{\boldsymbol{\rho}}(\boldsymbol{r}) = -\kappa^2 \boldsymbol{\nabla}(s_{e,\Gamma} * (\chi_\Sigma\,\phi_\Sigma))(\boldsymbol{r}), \quad \text{in } \mathbb{R}^3\,,$$

with a "suitable" mollifier, i.e., a mollifier which averages the discontinuity on Γ of the considered molecule and lets the remaining parts on Σ and Ω untouched.

The source terms, $\boldsymbol{\rho}$ and $\tilde{\boldsymbol{\rho}}$, differ only around Γ, where $\tilde{\boldsymbol{\rho}}$ is smooth and its derivatives exist. Thus, the gradient $\boldsymbol{\nabla}$ can be extracted from the volume integral without the appearance of the boundary integral.

$$\tilde{\boldsymbol{F}}(\boldsymbol{r}) \;=\; -\kappa^2 \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\,\boldsymbol{\nabla}_{r'}\big((s_{e,\Gamma} * (\chi_\Sigma\,\phi_\Sigma))(\boldsymbol{r}')\big)$$

$$\begin{aligned}
\Leftrightarrow \quad \tilde{\boldsymbol{F}}(\boldsymbol{r}) \;&=\; \underbrace{-\kappa^2 \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' \boldsymbol{\nabla}_{r'}\left[G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\left((s_{e,\Gamma} * (\chi_\Sigma\,\phi_\Sigma))(\boldsymbol{r}')\right)\right]}_{=0} \\
&\quad +\kappa^2 \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' \left(\boldsymbol{\nabla}_{r'} G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\right)\left((s_{e,\Gamma} * (\chi_\Sigma\,\phi_\Sigma))(\boldsymbol{r}')\right) \\
&=\; -\kappa^2 \boldsymbol{\nabla} \underbrace{\left(\int\limits_{\mathbb{R}^3} d\boldsymbol{r}' G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\left((s_{e,\Gamma} * (\chi_\Sigma\,\phi_\Sigma))(\boldsymbol{r}')\right)\right)}_{\tilde{f}_{\mathrm{N}}(\boldsymbol{r})}, \qquad \boldsymbol{r} \in \mathbb{R}^3
\end{aligned}$$

An application of the Yukawa operator $\mathcal{L}_\kappa$ yields

$$\mathcal{L}_\kappa f_{\mathrm{N}} = -\kappa^2 \left(s_{e,\Gamma} * (\chi_\Sigma\,\phi_\Sigma)\right), \quad \text{in } \mathbb{R}^3\,.$$

Up to now, we did not make any restriction on the parameter e. With the choice of a very small e, the situation resembles the one we started with closely. The limit $e \to 0$ yields the following Newton Scalar Model, which is abbreviated by NSM in the remainder of this work.

Theorem 5.2.1 *Newton Scalar Model (NSM): We define the following Newton potential*

$$\boldsymbol{F} = -\boldsymbol{\nabla} f_N = -\kappa^2 \boldsymbol{\nabla} \int\limits_{\mathbb{R}^3} d\boldsymbol{r}' G_{\mathcal{L}_\kappa}\,\chi_\Sigma \phi(\boldsymbol{r}'), \quad \boldsymbol{r} \in \mathbb{R}^3 \tag{5.16}$$

Inserting $\boldsymbol{F}$ into the nonlocal material equations, system (4.1) *on p. 51, and using the differential*

analog to Eq. (5.16) *given in Section 4.2, the Newton Scalar Model reads:*

$$\begin{aligned}
\varepsilon_\Omega \triangle \phi_\Omega &= -\rho, && \text{in } \Omega \\
\triangle(\varepsilon_\infty \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty) f_{N,\Sigma}) &= 0, && \text{in } \Sigma \\
\phi_\Omega &= \phi_\Sigma, && \text{in } \Gamma \\
\varepsilon_\Omega \partial_n \phi_\Omega &= \varepsilon_\infty \partial_n \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty)\, \partial_n f_{N,\Sigma}, && \text{in } \Gamma
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_\kappa f_{N,\Omega} &= 0, && \text{in } \Omega \\
\mathcal{L}_\kappa f_{N,\Sigma} &= -\kappa^2 \phi_\Sigma, && \text{in } \Sigma \\
f_{N,\Omega} &= f_{N,\Sigma}, && \text{on } \Gamma \\
\partial_n f_{N,\Omega} &= \partial_n f_{N,\Sigma}, && \text{on } \Gamma
\end{aligned}$$

In fact, in the NSM, the nonlocal dielectric function acts on the electrostatic potential instead on the electrostatic field. From the physical point of view, this means that the water molecules' reaction primarily depends on their potential energy in the electrostatic field and not – as it is the fundamental assumption of a mean field approach to account for the average polarization of the medium – on the electrostatic force acting on the medium's molecules:

$$\boldsymbol{D} = -\int_\Sigma d\boldsymbol{r}' \varepsilon(\boldsymbol{r} - \boldsymbol{r}') \boldsymbol{\nabla}_{r'} \phi(\boldsymbol{r}') \neq -\boldsymbol{\nabla} \int_\Sigma d\boldsymbol{r}' \varepsilon(\boldsymbol{r} - \boldsymbol{r}') \phi(\boldsymbol{r}')\,,$$

where the electrostatic field $\boldsymbol{E} = -\boldsymbol{\nabla}\phi$.

However, we have seen that a smoothing process which *is* a reasonable approximation – in particular, when considering that the boundaries of molecules is most probably smooth – yields this gradient model. Further, a comparison of the NVM and the NSM in Chapter 7 reveals that the differences are small.

5.3 Discussion and comparison

Before we apply the nonlocal scalar models on non-trivial geometries, we discuss their qualitative differences and the differences to the NVM and the local model (LM) in the case of spherical symmetry. For the sake of completeness, we therefore introduce the local model, which is based on the constant macroscopic dielectric functions ε_Ω and ε_Σ in Ω and Σ, respectively.

Theorem 5.3.1 *Local Model (LM)*

$$\begin{aligned}
\varepsilon_\Omega \triangle \phi_\Omega &= -\rho, && \text{in } \Omega \\
\varepsilon_\Sigma \triangle \phi_\Sigma &= 0, && \text{in } \Sigma \\
\varepsilon_\Omega \partial_n \phi_\Omega &= \varepsilon_\Sigma \partial_n \phi_\Sigma, && \text{on } \Gamma \\
\phi_\Omega &= \phi_\Sigma, && \text{on } \Gamma
\end{aligned}$$

The general shape of the solution for the Born sphere of all the electrostatic models can be expressed by the electrostatic potential ϕ and the correlation field potential f, which is one of

$\{F_{\mathrm{N}}, f_{\mathrm{N}}, F_{\mathrm{D},\Sigma}\}$:

$$\begin{aligned}
\phi_\Omega &= \frac{1}{4\pi}\Big[\frac{1}{\varepsilon_\Omega\, r} + \Big(\frac{1}{\varepsilon_\Sigma} - \frac{1}{\varepsilon_\Omega}\Big)\frac{1}{a} + \frac{B}{a}\Big] \\
f_\Omega &= \frac{1}{4\pi r}\, B_2\, \sinh(\kappa\, r) \\
\phi_\Sigma &= \frac{1}{4\pi r}\Big[\frac{1}{\varepsilon_\Sigma} + B\, e^{-\frac{1}{\lambda}(r-a)}\Big] \\
f_\Sigma &= \frac{1}{4\pi r}\Big[\frac{1}{\varepsilon_\Sigma} - B\, e^{-\frac{1}{\lambda}(r-a)}\frac{\varepsilon_\infty}{\varepsilon_\Sigma - \varepsilon_\infty}\Big]
\end{aligned}$$

Here, we assumed the Born sphere of a point charge with $q=1$ and radius a located in the origin of the coordinate system. The electrostatic potential ϕ_Σ is characterized by a $\frac{1}{r}$-part (Laplace-like) and a $\frac{e^{-r/\lambda}}{r}$-part (Yukawa-like). This corresponds to the fact that the differential equations in all models coincide in representing, on the one hand, the reaction of the point charge and, on the other hand, the reaction of the water network. Only the transmission and boundary conditions on Γ differ from each other, and this results in model dependent constants, B and B_2, which are listed in Tab. 5.1. We additionally listed the nonlocal point charge solution, which of course only exists for $\Sigma = \mathbb{R}^3$.

Tab. 5.1: Constants in the solution for the Born model of the different scalar models with $\kappa = \frac{1}{\lambda}\sqrt{\frac{\varepsilon_\infty}{\varepsilon_\Sigma}}$.

	B	B_2
NVM	$(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma})\,\frac{1}{1+a/\lambda+\frac{\varepsilon_\Sigma}{\varepsilon_\infty}(\frac{\cosh(\kappa\, a)\,\kappa\, a}{\sinh(\kappa\, a)}-1)}$	$B\,\frac{1}{\sinh(\kappa a)}\,\frac{\varepsilon_\Sigma}{\varepsilon_\infty-\varepsilon_\Sigma}$
DVM/DSM	$(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma})\frac{1}{1+a/\lambda}$	0
NSM	$(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma})\,\frac{1}{1+\frac{\sinh(\kappa a)}{\cosh(\kappa a)}\sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}}$	$B\,\frac{1}{\cosh(\kappa a)}\,\frac{\varepsilon_\infty}{\varepsilon_\Sigma-\varepsilon_\infty}\sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}$
point charge	$(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma})$	0
LM	0	0

With Tab. 5.1, we see that the complexity of the models - measured in terms of involved parameters $(\varepsilon_\Sigma, \varepsilon_\infty, \lambda, a)$ - decreases from the NVM to the LM. The constant B in all the nonlocal models has a linear dependency on $(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma})$, which is due to the high frequency constraint.

Fig. 5.6 shows the dependence of the parameter B on the ion radius a and the correlation length λ. In addition to the DSM (green) and the NSM (red), we plotted the point charge (blue), the Born solution of the NVM (yellow) and the Born solution for the local setting (black). From Fig. 5.6 we see that the point charge model and the local model correspond to the maximal and minimal nonlocal correlation effect, respectively. The DSM, the NSM, and the NVM are between these two extrema and therefore reveal an intermediate dielectric reaction. Only on the boundaries of the parameter regime, e.g., for $\lambda \to 0$ and $a \gg 1$, they differ significantly from each other, whereas in the intermediate region, they exhibit the same dependence on a and λ, namely a linear dependence on $\frac{1}{1+a/\lambda}$. When $\frac{a}{\lambda} \ll 1$, the water network does not "see" the dielectric boundary and therefore it does not "see" a disruptive factor of the correlations. The length scale on which the network is mediated is so large that the correlations are in fact not disturbed. In contrast, $\frac{a}{\lambda} \gg 1$ implies that the water network is built on a fractional amount of the extension of the spherical molecule. The water molecules "feel" the surface of the molecule, which is too extended for the correlations to still be sustained.

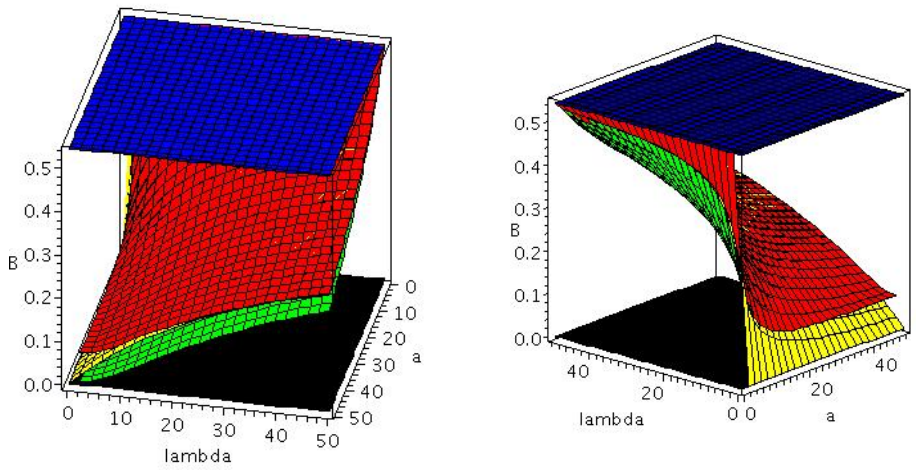


Fig. 5.6: Characteristic constant B of the Born model as a function of the sphere radius a and the correlation length λ; Point charge - blue; NSM - red; NVM - yellow; DSM/DVM - green; LM - black.

For the various models, we calculate the solvation free energy as a measure to assess the approximations we made in the scalar models. In order to evaluate the solvation energies, we process as described for the vector model in Section 3.3.3.4. The results are listed in Tab. 5.2.

Tab. 5.2: Solvation free energy for monoatomic ions (Born model); $\kappa = 1/23$ Å^{-1}, $\varepsilon_\Omega = 1\,\varepsilon_0$, $\varepsilon_\Sigma = 78.0\,\varepsilon_0$, $\varepsilon_\infty = 1.8\,\varepsilon_0$.

		Solvation energy $[KJ/mol]$					
monoatomic ion		exp.	local model	nonlocal models			local model
	R [Å]	data		NVM	DSM/DVM	NSM	$\varepsilon_\Sigma = \varepsilon_\infty$
Li^+	0.65	-511	-1055.03	-570.68	-565.97	-565.95	-474.99
Na^+	1.01	-411	-678.98	-395.52	-389.39	-389.35	-305.68
K^+	1.37	-337	-500.56	-309.89	-302.87	-302.81	-225.36
Rb^+	1.51	-316	-454.15	-287.07	-279.81	-279.73	-204.46
Cs^+	1.72	-284	-398.70	-259.35	-251.81	-251.72	-179.50
Mg^{2+}	0.62	-1906	-4424.33	-2376.74	-2358.48	-2358.41	-1991.90
Ca^{2+}	1.02	-1593	-2689.29	-1569.48	-1544.86	-1544.69	-1210.76
Sr^{2+}	1.2	-1447	-2285.90	-1377.03	-1350.43	-1350.22	-1029.15
Ba^{2+}	1.39	-1318	-1973.44	-1225.48	-1197.26	-1196.99	-888.47

The comparison confirms what we learned from the discussion before, namely that all the nonlocal models do a good job: the electrostatic part of the solvation free energy differs only slightly in the nonlocal models and has an excellent agreement with the experimental data. In contrast, the local model, which does not consider the water network, predicts solvation energies up to an order too large. This is reasonable, for the energy gain, when moving from a constant low dielectric to a constant high dielectric - as in the local model - is larger compared to the nonlocal model. In the nonlocal case the dielectric function varies, in fact weakens the response due to the water network.

In particular, the analysis performed in the last sections revealed that on Γ, the dielectric response of the nonlocal models (DSM, NSM, DVM, NVM) is dominated by the electronic polarization characterized by $\varepsilon_\infty < \varepsilon_\Sigma$. This is also shown in the last column of Tab. 5.2, where we listed the electrostatic contribution to the solvation free energy for a local model which has a solvent with macroscopic dielectric function ε_∞. The difference between the solvation energies of this local model and the nonlocal models can be interpreted as the amount of energy to regain the macroscopic response ε_Σ in the bulk.

In summary, we can say that although the two proposed scalar models have different approximations, they exhibit a very similar, physically reasonable behavior on spherical systems in comparison to the NVM which is historically the very first nonlocal model. Whether the potential description suffices to describe the electrostatics of non-trivial geometries is the focus of the following sections.

Chapter 6

Numerical solvers for nonlocal electrostatic models

In the preceding chapter we have developed two scalar models for nonlocal electrostatics. The application on spherically symmetric systems revealed that the approximations made in these scalar models do not change the main features of the Lorentzian water model. Now, we want to apply the scalar models to non-trivial geometries in order to further compare the different approximations and finally to propose a nonlocal model which we can optimize for a broad application. Once we allow for complex transmission regions Γ, we have to develop appropriate numerical schemes to solve for the linear, elliptic partial differential equations, which separately defined in Ω and Σ.

There are classical methods for solving linear, elliptic PDES such as

1. the Boundary Element Method (BEM):
   This method exploits that, if a fundamental solution of the differential operator exists, linear, differential equations can be rewritten into a set of boundary integral equations. These boundary integral equations describe the behavior of the fields on the transmission region Γ, which is the molecular surface in the biomolecular setting. The BEM uses a polygonal representation of the molecular surface, usually a triangular mesh, and solves for the complete jump conditions of the unknown fields at the surface (Dirichlet and Neumann trace). From these boundary data the solution in Ω and Σ can be retrieved via the representation formulas.

2. the Finite Element Method (FEM) and the Finite Difference Method (FDM):
   Both methods rely on a discretization of the 3D space to project the continuous solution on a finite set of basis functions [83, 114]. A discretized version of the differential equations is derived at each vertex, where the differential operators are transformed into differences involving the vertex and its neighbors on the grid. The FDM uses a structured, often Cartesian grid and this can limit its accuracy in critical regions, such as the surface of the molecule. In contrast, the FE grids are unstructured, which opens the possibility for local refinements.

All three approaches reduce the partial differential equation system (PDES) into a set of discrete algebraic equations, which need to be solved numerically. Most solvers use iterative schemes, such as – for linear systems – the stationary (Jacobi, Gauss-Seidel, Successive Over-Relaxation) and the non-stationary (Conjugate Gradient, Generalized Minimal Residual) methods or they use a direct solver, such as an LU or a Cholesky decomposition. As the concepts of the solving procedure are different for the BEM, FEM, and FDM, they have different advantages and disadvantages. These have to be taken into account for the final choice of the numerical scheme (see for an overview [52, 83, 114]).

Since our model equations are linear and elliptic, it is possible to solve them by the BEM. In fact, there exists a BEM solver for the local model (LM, Theorem 5.3.1), for the Dirichlet Scalar Model (DSM, Theorem 5.1.1) as well as for the Newton Vector Model (NVM, Theorem 4.3.2),

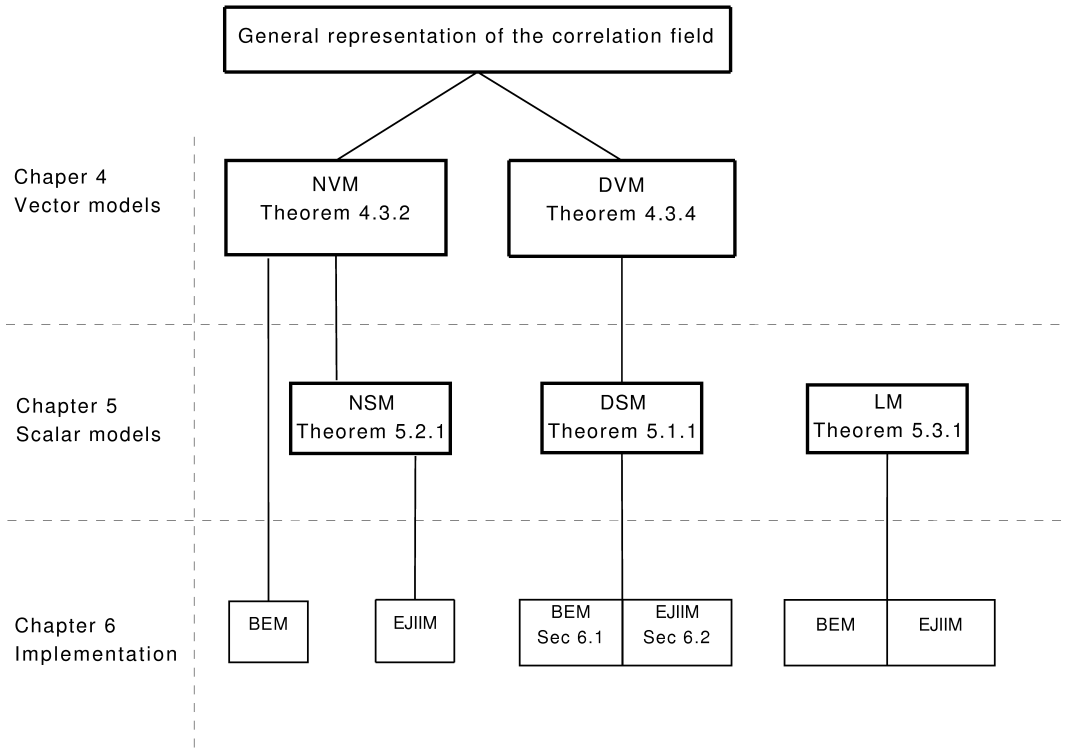


Fig. 6.1: Overview on the different models and developed/existing numerical solvers.

see [52] and [39], respectively. However, the use of the BEM imposes a serious restriction: the need for a fundamental solution of the differential operator. Due to this requirement, an extension or modification of the nonlocal PDES, such as the exchange of the dielectric operator or the inclusion of (non)linear counter-ion terms to account for ionic effects, becomes a challenging task.

FDMs for solving (nonlinear) PDES are widely used, as they are robust to implement and easily extensible. Their convergence behavior and accuracy is sufficient for most problems. This explains the high number of finite difference solvers freely available for electrostatic problems [9, 11, 85, 101, 129, 146] and the use of FDMs in various biomolecular software packages, such as BALL, H++, and PyMol. The accuracy of FDMs depends in part on an accurate surface description [31]. Especially if there are jump conditions – as it is the case for the Maxwell equations in the presence of dielectric discontinuities – algorithms that allow for a correct surface description are essential.

A possibility which combines the methodology of the simple finite difference discretization with a correct surface description and therefore acquires highly accurate results, is offered by the Explicit Jump Immersed Interface Method (EJIIM) together with suitable conditions along the artificial exterior boundary [79, 104, 143]. In this work, we develop an EJIIM for the scalar (non)local models (Theorems 5.1.1, 5.2.1 and 5.3.1). By using its flexibility, we will further extend this framework in Section 8.2 to account for the linearized effect of monovalent ions additionally solved in water.

The diagram in Fig. 6.1 gives an overview of all the models of the previous chapters. The last row identifies the existing BEM and the EJIIM solvers we develop. As for the DSM, a BEM implementation already exists, the first step towards an efficient finite difference solver was an analogous EJIIM implementation.

For the numerical solution process the differential system as it is given in Theorem 5.1.1 for the DSM is slightly rewritten: we consider a molecule $\mathcal{M}$ to be a set of N spheres at fixed positions $\boldsymbol{r}_i$ with radii R_i, and fixed charges q_i, $i = 1, \ldots, N$ at its centers. The sum over all point charges

forms the charge density

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N} q_i \, \delta(\boldsymbol{r} - \boldsymbol{r}_i) \,, \quad \boldsymbol{r} \in \mathbb{R}^3 \,. \tag{6.1}$$

Both, the BEM and the EJIIM, explicitly handle arbitrary transmission conditions on a smooth surface Γ in $\mathbb{R}^3$ and that is why we can use the decomposition of the electrostatic potential ϕ_Ω proposed in Section 3.1.4.1,

$$\phi_\Omega = \phi_\Omega^{reac} + \phi_{\text{mol}} \,.$$

The field ϕ_{mol} originates from all fixed charges q_i, treating them as if they were solvated in a medium of constant response ε_Ω. As ϕ_{mol} captures the singularities of the electric potential ϕ_Ω at the charge positions $\boldsymbol{r}_i$, such a decomposition is desirable in numerical applications to avoid numerical instabilities [1]. With Eq. (6.1) and the radiation condition, $\lim_{|\boldsymbol{r}|\to\infty} \phi_{\text{mol}} = 0$, we have

$$\triangle \phi_{\text{mol}}(\boldsymbol{r}) = -\frac{1}{\varepsilon_\Omega} \rho(\boldsymbol{r}) \quad \Rightarrow \quad \phi_{\text{mol}}(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_\Omega} \sum_{i}^{N} \frac{q_i}{|\boldsymbol{r} - \boldsymbol{r}_i|} \,, \quad \boldsymbol{r} \in \mathbb{R}^3 \,.$$

Therefore, the unknown part of ϕ_Ω reduces to ϕ_Ω^{reac}. As a second modification we do not consider the correlation field itself but the linear combination

$$\psi_\Sigma := (\varepsilon_\infty \, \phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty) \, f_\Sigma) \,.$$

The *dielectric potential* ψ_Σ fulfills the homogeneous Laplace equation and this allows, in the case of the EJIIM, to use a Fast Fourier Transformation (FFT) in the solution process.

The model presented in Theorem 5.1.1 expressed by the unknowns $\phi_\Omega^{reac}, \phi_\Sigma$, and ψ_Σ then reads:

$$\begin{aligned}
\Delta \phi_\Omega^{reac} &= 0, && \text{in } \Omega && \text{(6.2a)} \\
\Delta \psi_\Sigma &= 0, && \text{in } \Sigma && \text{(6.2b)} \\
\phi_\Sigma - \phi_\Omega^{reac} &= \phi_{\text{mol}}, && \text{on } \Gamma && \text{(6.2c)} \\
\partial_n \psi_\Sigma - \varepsilon_\Omega \partial_n \phi_\Omega^{reac} &= \varepsilon_\Omega \partial_n \phi_{\text{mol}}, && \text{on } \Gamma && \text{(6.2d)} \\
(\Delta - 1/\lambda^2) \, \phi_\Sigma + \frac{1}{\varepsilon_\Sigma \lambda^2} \, \psi_\Sigma &= 0, && \text{in } \Sigma && \text{(6.2e)} \\
\varepsilon_\infty \partial_n \phi_\Sigma - \partial_n \psi_\Sigma &= 0, && \text{on } \Gamma && \text{(6.2f)}
\end{aligned}$$

Using the DSM (system (6.2)), we give a short introduction to the BEM and the EJIIM in Sections 6.1 and 6.2, respectively. In Section 6.3 we explain the input data generation for the EJIIM, i.e., the accurate calculation of the intersections of the biomolecule's surface with a 3D Cartesian grid.

6.1 Boundary Element Method for the Dirichlet Scalar Model

The crucial idea of the BEM is (a) to derive an integral representation of the PDES and (b) based on the integral representation to deduce boundary integral equations whose solution is the full set of boundary values, i.e., the Dirichlet and the Neumann trace, see Definition 4.1. This set of boundary data is known as *Cauchy data* [121].

[1] In classical finite difference methods, the surface Γ only implicitly enters the differential equations and thus they are not designed for such a decomposition.

In Sections 6.1.1 and 6.1.2, we recall the representation formulas derived in [52] and the corresponding boundary integral equations for the DSM, respectively. Section 6.1.3 finally focuses on the input requirements for the BEM.

6.1.1 Representation formulas

In order to cast the DSM into a system of boundary integral equations, we first convert the PDES to its weak formulation by multiplying the differential equations with the fundamental solution of the respective differential operators [52, 121]. This has already been discussed in Section 4.3 for the Yukawa operator: shifting the differential operators from the unknown functions to the known fundamental solutions and applying Green's theorem introduces two boundary integrals and an additional domain integral over the inhomogeneity of the respective differential equation. The representation of the unknown field with these two boundary integrals and the volume integral is called representation formula.

For the system (6.2), the described procedure yields

$$\phi^{reac}(\boldsymbol{r}) = +\left(\tilde{V}^{\mathrm{L}}(\gamma_1^{int}\phi^{reac})\right)(\boldsymbol{r}) - \left(W^{\mathrm{L}}(\gamma_0^{int}\phi^{reac})\right)(\boldsymbol{r}), \qquad \boldsymbol{r}\in\Omega \tag{6.3}$$

$$\psi_\Sigma(\boldsymbol{r}) = -\left(\tilde{V}^{\mathrm{L}}(\gamma_1^{ext}\psi_\Sigma)\right)(\boldsymbol{r}) + \left(W^{\mathrm{L}}(\gamma_0^{ext}\psi_\Sigma)\right)(\boldsymbol{r}), \qquad \boldsymbol{r}\in\Sigma \tag{6.4}$$

$$\begin{aligned}\phi_\Sigma(\boldsymbol{r}) = &-\left(\tilde{V}^{\mathrm{Y}}(\gamma_1^{ext}\phi_\Sigma)\right)(\boldsymbol{r}) + \left(W^{\mathrm{Y}}(\gamma_0^{ext}\phi_\Sigma)\right)(\boldsymbol{r})\\ &+\left(\tilde{N}_0^{\mathrm{Y}}(-\frac{\kappa^2}{\varepsilon_\infty}\psi_\Sigma)\right)(\boldsymbol{r}), \qquad \boldsymbol{r}\in\Sigma\,,\end{aligned} \tag{6.5}$$

where we used the boundary integral operators defined in Definitions 4.2, 4.3, and 4.4 on p. 55. Eqs. (6.3) and (6.4) are similar to the well known representation formulas of the local model [52]. However, in Eq. (6.5) a Newton potential appears which has to be eliminated to take advantage of a pure boundary value representation. In the case of system (6.2), this domain integral can fortunately be reduced to boundary integrals using a dual reciprocity scheme, where the fundamental solution of the Yukawa equation is represented as the Laplacian of a function that can be easily determined [52]. This procedure yields

$$\begin{aligned}\phi_\Sigma(\boldsymbol{r}) = &-\left(\tilde{V}^{\mathrm{Y}}(\gamma_1^{ext}\phi_\Sigma)\right)(\boldsymbol{r}) + \left(W^{\mathrm{Y}}(\gamma_0^{ext}\phi_\Sigma)\right)(\boldsymbol{r})\\ &+\frac{1}{\varepsilon_\Sigma}\left\{\left((\tilde{V}^{\mathrm{Y}}-\tilde{V}^{\mathrm{L}})(\gamma_1^{ext}\psi_\Sigma)\right)(\boldsymbol{r}) - \left((W^{\mathrm{Y}}-W^{\mathrm{L}})(\gamma_0^{ext}\psi_\Sigma)\right)(\boldsymbol{r})\right\}.\end{aligned} \tag{6.6}$$

6.1.2 Boundary integral equations

In order to apply formulas (6.3), (6.4), and (6.6) for actual computations, we first need to determine the values of the relevant unknown potentials and their normal derivatives at the boundary, i.e., the Cauchy data of this problem. Using the transmission and boundary conditions in system (6.2), the set of unknown Cauchy data is given by $\gamma_0^{\mathrm{int}}\phi^{reac}$, $\gamma_1^{\mathrm{int}}\phi^{reac}$, and $\gamma_0^{\mathrm{int}}\psi$. They are calculated by carefully computing a limiting process for the evaluation of the representation formulas and their normal derivatives at the molecular surface. The computation of these limits is quite involved [52], and we only want to state the result here. To this end, we define the following boundary integral operators: the Dirichlet trace operators of the Single and Double layer potentials of the Laplace operator, $V^{\mathrm{L}}w$ and $K^{\mathrm{L}}w$, respectively, and the Dirichlet trace operators of the Single and Double

layer potentials of the Yukawa operator, $V^{\mathrm{Y}}w$ and $K^{\mathrm{Y}}w$:

$$\begin{aligned}
\big(V^{\mathrm{L}}w\big)(\boldsymbol{r}) &:= -\gamma_{0,r}^{\mathrm{int}}\,(\tilde{V}^{\mathrm{L}}w) = -\gamma_{0,r}^{\mathrm{int}}\int_{\Gamma}\big(\gamma_{0,r'}G_{\triangle}(\boldsymbol{r}-\boldsymbol{r}')\big)w(\boldsymbol{r}')\,d\Gamma_{r'}\\
\big(K^{\mathrm{L}}v\big)(\boldsymbol{r}) &:= -\gamma_{0,r}^{\mathrm{int}}\,(W^{\mathrm{L}}v) = -\gamma_{0,r}^{\mathrm{int}}\int_{\Gamma}\big(\gamma_{1,r'}G_{\triangle}(\boldsymbol{r}-\boldsymbol{r}')(\boldsymbol{r},\boldsymbol{r}')\big)w(\boldsymbol{r}')\,d\Gamma_{r'}\\
\big(V^{\mathrm{Y}}w\big)(\boldsymbol{r}) &:= -\gamma_{0,r}^{\mathrm{ext}}\,(\tilde{V}^{\mathrm{Y}}w) = -\gamma_{0,r}^{\mathrm{ext}}\int_{\Gamma}\big(\gamma_{0,r'}G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\big)v(\boldsymbol{r}')\,d\Gamma_{r'}\\
\big(K^{\mathrm{Y}}v\big)(\boldsymbol{r}) &:= -\gamma_{0,r}^{\mathrm{ext}}\,(W^{\mathrm{Y}}v) = -\gamma_{0,r}^{\mathrm{ext}}\int_{\Gamma}\big(\gamma_{1,r'}G_{\mathcal{L}_\kappa}(\boldsymbol{r}-\boldsymbol{r}')\big)v(\boldsymbol{r}')\,d\Gamma_{r'}\,,
\end{aligned}$$

where $\boldsymbol{r}\in\Gamma$, $G_{\triangle}$ denotes the fundamental solution of the Laplacian and $G_{\mathcal{L}_\kappa}$ the fundamental solution of the Yukawa operator. Furthermore, we introduce the geometrical quantity

$$\sigma(\boldsymbol{r}) := \lim_{\epsilon\to 0}\frac{1}{4\pi}\frac{1}{\epsilon^2}\int\limits_{\boldsymbol{r}'\in\Omega:|\boldsymbol{r}-\boldsymbol{r}'|=\epsilon} d\Gamma_{r'}\,,$$

and

$$\beta := -\left(1-\sigma-K^{\mathrm{Y}}+\frac{\varepsilon_\Omega}{\varepsilon_\Sigma}\left(K^{\mathrm{Y}}-K^{\mathrm{L}}\right)\right)\left(\gamma_0^{\mathrm{int}}\phi_{\mathrm{mol}}\right)-\left(\frac{\varepsilon_\Omega}{\varepsilon_\infty}V^{\mathrm{Y}}-\frac{\varepsilon_\Omega}{\varepsilon_\Sigma}\left(V^{\mathrm{Y}}-V^{\mathrm{L}}\right)\right)\left(\gamma_1^{\mathrm{int}}\phi_{\mathrm{mol}}\right)\,.$$

With

$$\eta := \frac{1}{\varepsilon_\infty}\left\{\psi_\Sigma-\varepsilon_\Omega\phi_{\mathrm{mol}}\right\}\,,$$

the system of integral equations then reads

$$\begin{pmatrix}(1-\sigma)I-K^{\mathrm{Y}} & \frac{\varepsilon_\Omega}{\varepsilon_\infty}V^{\mathrm{Y}}-\frac{\varepsilon_\Omega}{\varepsilon_\Sigma}\left(V^{\mathrm{Y}}-V^{\mathrm{L}}\right) & \frac{\varepsilon_\infty}{\varepsilon_\Sigma}\left(K^{\mathrm{Y}}-K^{\mathrm{L}}\right)\\ \sigma I+K^{\mathrm{L}} & -V^{\mathrm{L}} & 0\\ 0 & \frac{\varepsilon_\Omega}{\varepsilon_\infty}V^{\mathrm{L}} & (1-\sigma)I-K^{\mathrm{L}}\end{pmatrix}\begin{pmatrix}\gamma_0^{\mathrm{int}}\phi^{reac}\\ \gamma_1^{\mathrm{int}}\phi^{reac}\\ \gamma_0^{\mathrm{ext}}\eta\end{pmatrix}=\begin{pmatrix}\beta\\0\\0\end{pmatrix}\,.$$

The potentials appearing in this system are approximated as low-rank polynomials defined on the triangles of an appropriate triangulation of the molecular surface. Denoting the resulting approximations of $\gamma_0^{\mathrm{int}}\phi^{reac}$, $\gamma_1^{\mathrm{int}}\phi^{reac}$, and $\gamma_0^{\mathrm{int}}\eta$ by $\hat{\boldsymbol{u}},\hat{\boldsymbol{q}}$, and $\hat{\boldsymbol{w}}$, respectively, the system reduces to the discrete algebraic form

$$\begin{pmatrix}(1-\sigma)\underline{\underline{\mathbf{I}}}-\underline{\underline{\mathbf{K}}}^{\mathrm{Y}} & \frac{\varepsilon_\Omega}{\varepsilon_\infty}\underline{\underline{\mathbf{V}}}^{\mathrm{Y}}-\frac{\varepsilon_\Omega}{\varepsilon_\Sigma}\left(\underline{\underline{\mathbf{V}}}^{\mathrm{Y}}-\underline{\underline{\mathbf{V}}}^{\mathrm{L}}\right) & \frac{\varepsilon_\infty}{\varepsilon_\Sigma}\left(\underline{\underline{\mathbf{K}}}^{\mathrm{Y}}-\underline{\underline{\mathbf{K}}}^{\mathrm{L}}\right)\\ \sigma\underline{\underline{\mathbf{I}}}+\underline{\underline{\mathbf{K}}}^{\mathrm{L}} & -\underline{\underline{\mathbf{V}}}^{\mathrm{L}} & \underline{\underline{\mathbf{0}}}\\ \underline{\underline{\mathbf{0}}} & \frac{\varepsilon_\Omega}{\varepsilon_\infty}\underline{\underline{\mathbf{V}}}^{\mathrm{L}} & (1-\sigma)\underline{\underline{\mathbf{I}}}-\underline{\underline{\mathbf{K}}}^{\mathrm{L}}\end{pmatrix}\begin{pmatrix}\hat{\boldsymbol{u}}\\ \hat{\boldsymbol{q}}\\ \hat{\boldsymbol{w}}\end{pmatrix}=\begin{pmatrix}\beta\\0\\0\end{pmatrix}\,,$$

where $\underline{\underline{\mathbf{K}}}^{\mathrm{Y}},\underline{\underline{\mathbf{V}}}^{\mathrm{Y}},\underline{\underline{\mathbf{K}}}^{\mathrm{L}},\underline{\underline{\mathbf{I}}}$, and $\underline{\underline{\mathbf{V}}}^{\mathrm{L}}$ denote the $n\times n$-matrices resulting in the application of the corresponding operators to the polynomial basis of the approximation and $\underline{\underline{\mathbf{0}}}$ the $n\times n$ zero matrix. The integer n is assumed to be the number of triangles in the surface discretization. This result can be compared to the BEM system for local electrostatics [52]

$$\begin{aligned}\left\{\left(1+(\frac{\varepsilon_\Omega}{\varepsilon_\Sigma}-1)\sigma\right)\underline{\underline{\mathbf{I}}}+\left(\frac{\varepsilon_\Omega}{\varepsilon_\Sigma}-1\right)\underline{\underline{\mathbf{K}}}^{\mathrm{L}}\right\}\hat{\boldsymbol{u}} &= \left(\underline{\underline{\mathbf{K}}}^{\mathrm{L}}+(\sigma-1)\underline{\underline{\mathbf{I}}}\right)\hat{\boldsymbol{u}}_{\mathrm{mol}}-\frac{\varepsilon_\Omega}{\varepsilon_\Sigma}\underline{\underline{\mathbf{V}}}^{\mathrm{L}}\hat{\boldsymbol{q}}_{\mathrm{mol}}\\ \underline{\underline{\mathbf{V}}}^{\mathrm{L}}\hat{\boldsymbol{q}} &= (\sigma\underline{\underline{\mathbf{I}}}+\underline{\underline{\mathbf{K}}}^{\mathrm{L}})\hat{\boldsymbol{u}}\,.\end{aligned}$$

The current implementation of the BEM solver for the DSM in `C` uses the ATLAS library [139] for an efficient solution of the linear system of equations. With the complete set of Cauchy data, the electrostatic potential at the points of interest can be evaluated in a postprocessing step using the representation formulas, Eqs. (6.3), (6.4), and (6.6).

6.1.3 Surface triangulation

The required geometric input data for the BEM is a numerical approximation of the molecular surface Γ – typically, a triangulation of the surface. The generation of such triangulations has received considerable attention in the literature, and most surface definitions can today be triangulated with good quality (see, [21, 114, 148]). However, molecular surfaces tend to be very complex – much more complex than, e.g., typical machine parts in CAD applications – and a triangulation with guaranteed quality will often require tens or hundreds of millions of triangles. Thus, input meshes for the BEM are often subjected to mesh coarsening algorithms.

The triangulations of the small molecules used in Section 6.4.2 are generated with the freely available mesh generator of Cheng and Edelsbrunner [22]. Surface representations of the biomolecules used in Section 7.2 are derived from exact intersection points of the molecule in a 3-dimensional Cartesian grid as described in Section 6.3.

6.2 Explicit Jump Immersed Interface Method for the Dirichlet Scalar Model

The Explicit Jump Immersed Interface Method (EJIIM) is a finite difference method that works on an equidistant Cartesian grid. The method can handle non-grid aligned discontinuities in the PDES, because in the EJIIM, close to discontinuities the standard finite difference approximations are modified by correction terms which involve the jumps in the function and its derivatives.

In this section, we recapitulate the main idea of the method and develop the algebraic system of the DSM (system (6.2)). A more detailed introduction to EJIIM can be found in the literature [104, 105, 107, 113, 140, 144] as well as on the website [106] where V. Rutka provides a descriptive documentation of the EJIIM for the inner 2D Poisson boundary value problem.

After a few definitions we set up the algebraic equations for system (6.2) in Section 6.2.1 and complete the EJIIM system with the algebraic equations for the unknown jumps in Section 6.2.2. The solution process is discussed in Section 6.2.3. Section 6.2.4 finally describes an estimation for the boundary values of the electrostatic and the dielectric potential. Such an estimation is always required in FDMs.

The EJIIM implementation of the DSM has been developed by V. Rutka. This implementation has been used for all the calculations presented in this work. Further, it was the starting point for modifications which we incorporated to adapt the original code to describe the NSM and to account for ionic solvents.

6.2.1 EJIIM discretization

The system (6.2) is a *transmission* problem for the electrostatic potential ϕ and a *boundary value* problem for the dielectric potential ψ. To make a global differential description possible and thereby to take advantage of *global* solvers such as the FFT, the system (6.2) is extended by an additional equation,

$$\Delta\psi_\Omega = 0 \text{ in } \Omega\,, \quad \psi_\Omega = 0 \text{ on } \Gamma\,,$$

forcing ψ_Ω to be zero. Along the exterior boundary $\partial\Sigma := \partial(\Sigma \cup \Gamma \cup \Omega)$, Dirichlet boundary values are given as we will discuss in Section 6.2.4.

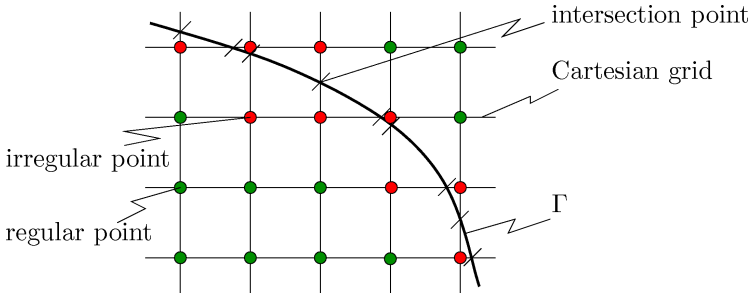


Fig. 6.2: The (red) grid points next to the intersections of the surface Γ with the regular grid are irregular grid points, the remaining grid points (green) are called regular.

Remark 6.1 Requiring ψ_Ω to be zero is different from the local electrostatic condition $\psi_\Omega = \varepsilon_\Omega \phi_\Omega$. However, forcing $\psi_\Omega = 0$ does not have any influence on the system (6.2) and the "physical" ψ_Ω can be easily obtained in a postprocessing step.

Up to now, Σ filled the whole $\mathbb{R}^3$, however, using finite difference methods we have to confine Σ to a finite volume, which is described by a discrete grid. In the following, we define the grid and further quantities required to formulate the differential equation system in the EJIIM:

Definition 6.1 Regular grid and intersections:
A regular grid with spacing h is imposed over a finite volume, i.e., over $\Omega \cup \Gamma \cup \Sigma$, leading to a total number of grid points $\{n_x, n_y, n_z\}$ in the three Cartesian directions. We will use the standard notation $u_{i,j,k} := u(x_i, y_j, z_k)$ for the quantity u at the grid point (x_i, y_j, z_k), $i \in \{1, \ldots, n_x\}$, $j \in \{1, \ldots, n_y\}$, $k \in \{1, \ldots, n_z\}$. Points where the interface Γ intersects the grid lines are called *intersections*.

The 2-dimensional regular Cartesian grid is illustrated in Fig. 6.2. The intersections of the surface Γ with this grid are indicated by small cross lines.

Definition 6.2 Standard central finite difference operator:
In the finite difference method the differential operators have to be discretized. The Laplace operator is the only differential operator appearing in system (6.2). Assuming a grid spacing h, we define the standard central finite difference operator

$$\partial^h_{xx} u(x_i, y_j, z_k) := \frac{u(x_{i-1}, y_j, z_k) - 2u(x_i, y_j, z_k) + u(x_{i+1}, y_j, z_k)}{h^2} .$$

$\partial^h_{xx} u(x_i, y_j, z_k)$ is the $\mathcal{O}(h^2)$-approximation of the second partial derivative of u at point (x_i, y_j, z_k) in x-direction,

$$\partial_{xx} u(x_i, y_j, z_k) = \partial^h_{xx} u(x_i, y_j, z_k) + \mathcal{O}(h^2) .$$

With an analogous finite difference formulation in y- and z-direction, we can finally express the discrete Laplace operator.

Definition 6.3 Regular and irregular points:
A grid point is classified as *regular*, if the standard central finite difference approximation of system (6.2) is not influenced by the interface Γ. All other grid points are called *irregular*.

In Fig. 6.2 the regular grid points are marked as green filled circles, whereas the red ones denote irregular grid points.

Definition 6.4 One sided values and jump variables:
For a piecewise continuous function $u : \mathbb{R}^3 \supset \Omega \cup \Gamma \cup \Sigma \to \mathbb{R}$ with a possible discontinuity across the interface Γ and $\alpha \in \Gamma$, we define *the one-sided values* of u as

$$u_\Omega(\alpha) := \lim_{\Omega \ni \boldsymbol{r} \to \alpha} u(\boldsymbol{r}) , \quad u_\Sigma(\alpha) := \lim_{\Sigma \ni \boldsymbol{r} \to \alpha} u(\boldsymbol{r})$$

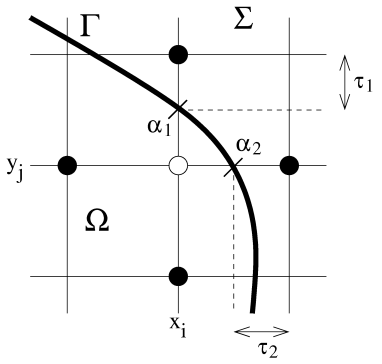


Fig. 6.3: Two dimensional illustration of a discretization at an irregular grid point (x_i, y_j) with two intersection points α_1 and α_2.

and the *jump* as

$$[u]_\alpha := u_\Sigma(\alpha) - u_\Omega(\alpha)\,.$$

Mostly, the index and the argument α will be omitted for the jump condition, meaning that the relation holds everywhere along Γ.

In order to discretize the equations, standard finite differences are used at all regular points. At all irregular points we introduce jump dependent *correction terms* to achieve an $\mathcal{O}(h)$ approximation. In a planar situation as shown in Fig. 6.3, we have

$$\partial_{xx}u(x_i,y_j) \approx \partial^h_{xx}u(x_i,y_j) - \frac{1}{h^2}\left([u]_{\alpha_2} + \tau_2[\partial_x u]_{\alpha_2} + \frac{\tau_2^2}{2}[\partial_{xx}u]_{\alpha_2}\right) \tag{6.7a}$$

$$\partial_{yy}u(x_i,y_j) \approx \underbrace{\partial^h_{yy}u(x_i,y_j)}_{\text{standard Laplace}} \underbrace{-\frac{1}{h^2}\left([u]_{\alpha_1} + \tau_1[\partial_y u]_{\alpha_1} + \frac{\tau_1^2}{2}[\partial_{yy}u]_{\alpha_1}\right)}_{\text{correction terms}}. \tag{6.7b}$$

The jumps in the Cartesian derivatives appearing in the correction terms in Eq. (6.7) are not explicitly known and therefore we introduce them as new *additional variables*. These so called jump variables J^ϕ and J^ψ account for the jumps in the potentials ϕ, ψ, and their derivatives. System (6.2) requires the following jump variables:

$$\begin{aligned} J^\phi &= ([\phi]_{\alpha_1}, [\partial_x\phi]_{\alpha_1}, [\partial_y\phi]_{\alpha_1}, [\partial_z\phi]_{\alpha_1}, [\partial_{xx}\phi]_{\alpha_1}, [\partial_{yy}\phi]_{\alpha_1}, [\partial_{zz}\phi]_{\alpha_1}, [\phi]_{\alpha_2}, ..., [\partial_{zz}\phi]_{\alpha_N})^T \\ J^\psi &= ([\psi]_{\alpha_1}, [\partial_x\psi]_{\alpha_1}, [\partial_y\psi]_{\alpha_1}, [\partial_z\psi]_{\alpha_1}, [\partial_{xx}\psi]_{\alpha_1}, [\partial_{yy}\psi]_{\alpha_1}, [\partial_{zz}\psi]_{\alpha_1}, [\psi]_{\alpha_2}, ..., [\partial_{zz}\psi]_{\alpha_N})^T \end{aligned}$$

In order to write the corrected approximations in a matrix-vector form, we introduce the discrete solution vectors Φ and Ψ containing values of ϕ, with $\phi_{|\Omega} := \phi^{reac}$ and $\phi_{|\Sigma} = \phi_\Sigma$, and ψ at the grid points as well as vectors F_1 and F_2 storing the right hand side of (6.2)(a,b,e). The discrete finite difference form of system (6.2) is then written as follows

$$\underbrace{\begin{pmatrix} \underline{\underline{\mathbf{\Delta^h}}} - 1/\lambda^2\,\underline{\underline{\mathbf{I}}}_\Sigma & \kappa^2/\varepsilon_\infty\,\underline{\underline{\mathbf{I}}}_\Sigma \\ \underline{\underline{\mathbf{0}}} & \underline{\underline{\mathbf{\Delta^h}}} \end{pmatrix}}_{=:\underline{\underline{\mathbf{A}}}} \underbrace{\begin{pmatrix}\Phi\\ \Psi\end{pmatrix}}_{=:W} + \underbrace{\begin{pmatrix} \underline{\underline{\mathbf{C_{11}}}} & \underline{\underline{\mathbf{0}}} \\ \underline{\underline{\mathbf{0}}} & \underline{\underline{\mathbf{C_{22}}}} \end{pmatrix}}_{=:\underline{\underline{\mathbf{C}}}} \underbrace{\begin{pmatrix}J^\phi\\ J^\psi\end{pmatrix}}_{=:J} = \underbrace{\begin{pmatrix}F_1\\ F_2\end{pmatrix}}_{=:F}, \tag{6.8}$$

where $\underline{\underline{\mathbf{\Delta^h}}}$ is the sparse discrete Laplace matrix (defined in $\Omega \cup \Gamma \cup \Sigma$) and $\underline{\underline{\mathbf{I}}}_\Sigma$ is a diagonal matrix with ones on the rows corresponding to the grid points in Σ and zeros otherwise. The term $\underline{\underline{\mathbf{C}}}\,J$

codes the correction terms at the irregular grid points as they are noted in Eqs. (6.7).

6.2.2 Approximating the jump variables

It remains to determine the unknown jumps J^ϕ and J^ψ. The equations for the jump variables are gained from the interface conditions in system (6.2), which are given in the local coordinates tangential and normal to the surface. In order to apply the method in Eq. (6.8), we need the jumps in Cartesian derivatives of the unknown functions ϕ and ψ:

- the 0th order jump for ϕ is given by (6.2c), $[\phi] = \phi_\Sigma - \phi^{reac} = \phi_{mol}$, and this also defines the jump in the tangential derivatives [29].
  For ψ, we set the trivial jump condition $[\psi] = \psi_\Sigma - \psi_\Omega = \psi_\Sigma$ as we need $\psi_\Omega = 0$.

- the 1st order jumps: let $\boldsymbol{t}$ and $\boldsymbol{s}$ be two tangents of Γ such that $(\boldsymbol{n}, \boldsymbol{t}, \boldsymbol{s})$ forms an orthonormal, positively oriented trihedral. Further, let $\underline{\underline{\mathbf{m}}} \in \mathbb{R}^{3\times 3}$ be the transformation matrix with columns given by vectors $\boldsymbol{n}$, $\boldsymbol{t}$, and $\boldsymbol{s}$. Then, considering system (6.2), we obtain

$$\begin{bmatrix} \partial_x\phi \\ \partial_y\phi \\ \partial_z\phi \end{bmatrix} = \underline{\underline{\mathbf{m}}} \begin{bmatrix} \partial_n\phi \\ \partial_t\phi \\ \partial_s\phi \end{bmatrix} = \underline{\underline{\mathbf{m}}} \begin{pmatrix} \frac{\varepsilon_\Omega - \varepsilon_\infty}{\varepsilon_\infty}\partial_n\phi^{reac} + \frac{\varepsilon_\Omega}{\varepsilon_\infty}\partial_n\phi_{mol} \\ \partial_t\phi_{mol} \\ \partial_s\phi_{mol} \end{pmatrix} .$$

  For ψ, we have $\psi_\Omega = 0$ and, combined with the boundary condition in system (6.2), we obtain

$$\begin{bmatrix} \partial_x\psi \\ \partial_y\psi \\ \partial_z\psi \end{bmatrix} = \begin{pmatrix} \partial_x\psi_\Sigma \\ \partial_y\psi_\Sigma \\ \partial_z\psi_\Sigma \end{pmatrix} = \underline{\underline{\mathbf{m}}} \begin{pmatrix} \partial_n\psi_\Sigma \\ \partial_t\psi_\Sigma \\ \partial_s\psi_\Sigma \end{pmatrix} = \underline{\underline{\mathbf{m}}} \begin{pmatrix} \frac{1}{2}\varepsilon_\Omega\partial_n\phi^{reac} + \frac{1}{2}\varepsilon_\Omega\partial_n\phi_{mol} + \frac{1}{2}\varepsilon_\infty\partial_n\phi_\Sigma \\ \partial_t\psi_\Sigma \\ \partial_s\psi_\Sigma \end{pmatrix} .$$

- all necessary second order jumps are obtained by extrapolation:

$$\left[\partial^2\phi\right] = \partial^2\phi_\Sigma - \partial^2\phi^{reac} \quad , \quad \left[\partial^2\psi\right] = \partial^2\psi_\Sigma - \partial^2\psi_\Omega = \partial^2\psi_\Sigma$$

We marked the known parts in the jump conditions in blue, whereas the unknown parts are marked in red. In fact, these are the one sided values of the solutions ϕ and ψ and their derivatives on Γ [2]. We denote them the *one sided extrapolations*. Since the finite difference method yields only a discretized version of the solutions, namely the grid solution vector W in Eq. (6.8), we have to extrapolate the grid values to the one sided extrapolations. With the restriction that the extrapolation is done via a linear combination of the grid point solutions [113], the transmission conditions basically form the conditional equations for the jumps

$$\underline{\underline{\mathbf{D}}} W + \underline{\underline{\mathbf{I}}} J = \tilde{F} . \tag{6.9}$$

Here, $\tilde{F}$ contains all known information about the jumps at the intersection points, i.e., the blue parts in the jump equations, such as ϕ_{mol} and $\boldsymbol{\nabla}\phi_{mol}$. In contrast, $\underline{\underline{\mathbf{D}}} W$ comprises the part of the jump conditions which depend on the solution vector W itself, i.e., the one sided extrapolations of the solutions and their derivatives, for example all terms depending on ψ_Σ in the above equations.

We still need a method to express the one sided extrapolations as a linear function of the discrete grid solutions. This means that we have to determine the matrix $\underline{\underline{\mathbf{D}}}$. To this end, we introduce for every intersection point two quadratic polynomials defined on each side of Γ [104]. These

[2] These unknowns are comparable with the Cauchy data in the BEM, see [142].

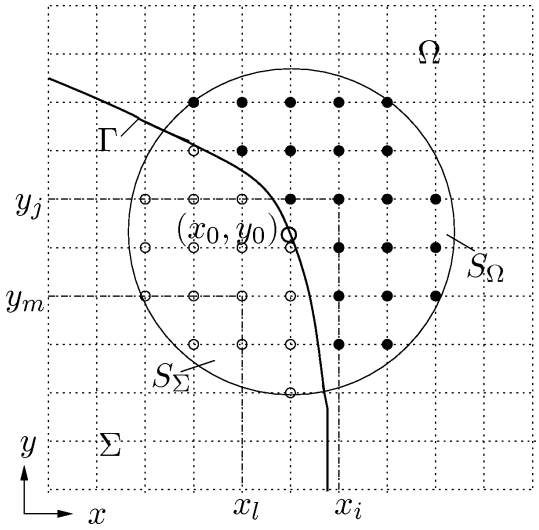


Fig. 6.4: Approximating the jump condition at the intersection point (x_0, y_0, z_0) by a least square fit of quadratic polynomials defined in the stencil S_Ω and S_Σ, respectively. The figure shows the xy-plane, i.e., a cross section through $z = z_0$.

polynomials transfer the transmission conditions, which have to be fulfilled on Γ, to the grid point solutions.

Fig. 6.4 illustrates the xy-plane of a 3D-stencil S around a considered intersection point (x_0, y_0, z_0) which determines the coefficients in the polynomial approach of, for example, ϕ:

$$\begin{aligned}
\phi_\Omega^{reac}(x,y,z) \;\approx\; p_\Omega(x,y,z) := \;& p_1 + p_2(x_0 - x) + p_3(y_0 - y) + p_4(z_0 - z) + \\
& p_5(x_0 - x)^2 + p_6(y_0 - y)^2 + p_7(z_0 - z)^2, \qquad \forall (x,y,z) \in S \subset \Omega \\
\phi_\Sigma(x,y,z) \;\approx\; p_\Sigma(x,y,z) := \;& g_1 + g_2(x_0 - x) + g_3(y_0 - y) + g_4(z_0 - z) + \\
& g_5(x_0 - x)^2 + g_6(y_0 - y)^2 + g_7(z_0 - z)^2, \qquad \forall (x,y,z) \in S \subset \Sigma
\end{aligned}$$

Now, $\underline{\underline{\mathbf{D}}}$ is determined by finding optimal sets $(p_i, i = \{1, .., 7\})$ and $(g_i, i = \{1, .., 7\})$ of coefficients fulfilling the following three conditions:

1. For every grid point (x_i, y_j, z_k) in the stencil S_Ω (black filled circles in Fig. 6.4) and every point (x_l, y_m, z_n) in the stencil S_Σ (white filled circles in Fig. 6.4), respectively, it holds

$$\begin{aligned}
\phi_\Omega^{reac}(x_i, y_j, z_k) &\approx p_\Omega(x_i, y_j, z_k) := p_1 + p_2 h_i + p_3 k_j + p_4 t_k + p_5 h_i^2 + p_6 s_j^2 + p_7 t_k^2 \\
\phi_\Sigma(x_l, y_m, z_n) &\approx p_\Sigma(x_l, y_m, z_n) := g_1 + g_2 u_l + g_3 v_m + g_4 w_n + g_5 u_l^2 + g_6 v_m^2 + g_7 w_n^2 \,,
\end{aligned}$$

   where $h_i := (x_0 - x_i)$, $s_j := (y_0 - y_j)$, $t_k := (z_0 - z_k)$, $u_l := (x_0 - x_l)$, $v_m := (y_0 - y_m)$, and $w_n := (z_0 - z_n)$. This is guaranteed by a weighted least square fit of the polynomial on the discrete grid solution in the respective region, for example in Ω

$$\sum_{(x_i,y_j,z_k)\in S_\Omega} w_{ijk}^2 \left(p^\Omega(x_i, y_j, z_k) - \phi_\Omega^{reac}(x_i, y_j, z_k)\right)^2 \to \min$$

$$\text{with } w_{ijk} = \left(1 + d(x_i, y_j, z_k)/h\right)^{-1}, \quad d(x_i, y_j, z_k) = \sqrt{h_i^2 + s_j^2 + t_k^2}\,.$$

   We used the least square fit described in [104, 113].

2. The difference between the polynomials p_Ω and p_Σ at every intersection point has to fulfill the jump conditions. For instance,

$$J^\phi = \begin{pmatrix} [\phi] \\ [\partial_x \phi] \\ \vdots \\ [\partial_{xx} \phi] \\ \vdots \end{pmatrix} \approx \begin{pmatrix} g_1 - p_1 \\ g_2 - p_2 \\ \vdots \\ g_5 - p_5 \\ \vdots \end{pmatrix}$$

3. The polynomial fit has to fulfill the differential equation. For instance in Ω, p_Ω has to fulfill

$$\triangle p_\Omega = 0 \qquad \Leftrightarrow \qquad p_5 + p_6 + p_7 \approx 0$$

The first and the second conditions determine the matrix $\underline{\underline{\mathbf{D}}}$, i.e., the optimal linear combination of the grid solution in the corresponding stencils to represent the one sided extrapolations. The third condition is not necessary but increases the accuracy of the method and reduces the iteration counts considerably [105]. This condition is guaranteed by adding a weighted penalty term in system (6.9) which corresponds to the side condition (similar to the Lagrange multiplier).

Remark 6.2 In the 3D case the stencils S_Ω and S_Σ contain too many points to take all of them into account for the minimization problem. In [104], V. Rutka proposed a random stencil selection in order to reduce the numerical costs. This random selection is also used in all presented calculations.

6.2.3 Solving the discrete system

With systems (6.8) and (6.9) we obtain the full EJIIM discretization (for clarity we now skip the subscript $\underline{\underline{*}}$):

$$\begin{pmatrix} A & C \\ D & I \end{pmatrix} \begin{pmatrix} W \\ J \end{pmatrix} = \begin{pmatrix} F \\ \tilde{F} \end{pmatrix} \tag{6.10}$$

System (6.10) is solved iteratively using a *stabilized Bi-Conjugate Gradient method* for non-symmetric matrices (BiCGSTAB) [66] using a Schur complement for the variable J:

$$(I - DA^{-1}C)J = \tilde{F} - DA^{-1}F\,. \tag{6.11}$$

Once the jump vector J has been found, W is obtained from $W = A^{-1}(F - CJ)$.

In each iteration, when solving (6.11), we need to apply the operator A^{-1}. To this end, the following discrete problem has to be solved

$$\begin{pmatrix} \Delta^h - \frac{1}{\lambda^2} I & \kappa^2/\varepsilon_\infty\, I \\ 0 & \Delta^h \end{pmatrix} \begin{pmatrix} \Phi \\ \Psi \end{pmatrix} = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix},$$

where Φ and Ψ are discrete solution vectors at the grid points and P_1, P_2 are some right hand side vectors, changing in each iteration. Solving is done in two steps:

1. $\Psi = (\Delta^h)^{-1} P_2$
2. $\Phi = (\Delta^h - \frac{1}{\lambda^2} I)^{-1}(P_1 - \frac{\kappa^2}{\varepsilon_\infty} I\Psi)$

The second equation is solved by an additional (inner) BiCGSTAB iteration. Iteration counts in this formulation are high (around 50-60 to reach 10^{-10} tolerance) and are reduced to typically 5-6

iterations by using a discrete Laplace as preconditioner:

$$(\Delta^h)^{-1}(\Delta^h - \frac{1}{\lambda^2} I)\Phi = (\Delta^h)^{-1}(P_1 - \frac{\kappa^2}{\varepsilon_\infty} I\Psi)\,.$$

The operator $(\Delta^h)^{-1}$ is always applied by an FFT-based solver [141].

6.2.4 Approximations of the exterior boundary condition

In contrast to the BEM, where the outer region does not have to be limited to a finite volume, this is necessary when solving the differential equations using a finite difference method. Thus, in addition to setting up the algebraic equations, we have to estimate the solution of the nonlocal electrostatic fields, for instance, ϕ_Σ and ψ_Σ in the DSM, on the exterior border of Σ, i.e., on $\partial\Sigma$. In the following, we propose different boundary conditions:

- homogeneous boundary condition:
  the simplest ansatz is to set all boundary values to zero

  $$\phi_\Sigma(\boldsymbol{r}) = 0 \quad \text{and} \quad \psi_\Sigma(\boldsymbol{r}) = 0, \quad \boldsymbol{r} \in \partial\Sigma\,.$$

  This is justifiable as the physical fields are restricted to the radiation condition introduced in Section 2.1.
- boundary condition without inner medium:
  we assume that the charge distribution is directly immersed in water. The resulting boundary value calculation is derived in Section 6.2.4.1 and is specified in Definition 6.6 on p. 98.
- boundary condition of an approximated charge distribution without inner medium:
  in addition to the assumption that the charge distribution is immersed in water we approximate the charge distribution by an effective dipole, see Definition 6.7 on p. 98.

All these proposed boundary value approximations can be combined with a *focusing method*, meaning to compute the solutions on a larger but coarse grid with one of the previously defined boundary conditions and to take its result for initializing the boundary values of the original (fine) grid.

6.2.4.1 Boundary condition without inner medium and simplified charge distribution

Although we have to cope with a complex dielectric response of the biomolecular system, which is reflected by the transmission problem of the electrostatic models, we assume that the behavior far away from the surface is mainly determined by the charge distribution and the dielectric response outside the molecule. Such a reduction of the real setting is shown in Fig. 6.5.

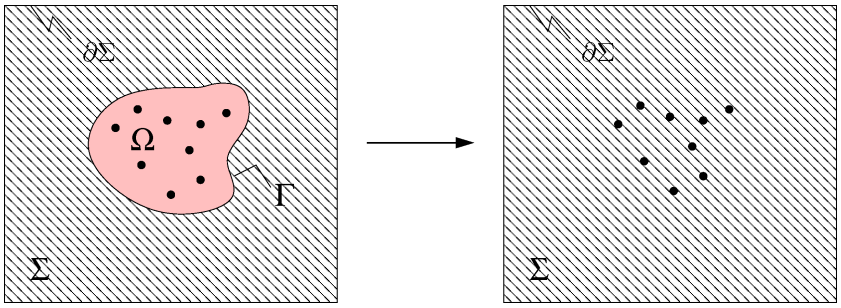


Fig. 6.5: Approximation of the potential values on $\partial\Sigma$ by neglect of the dielectric boundary.

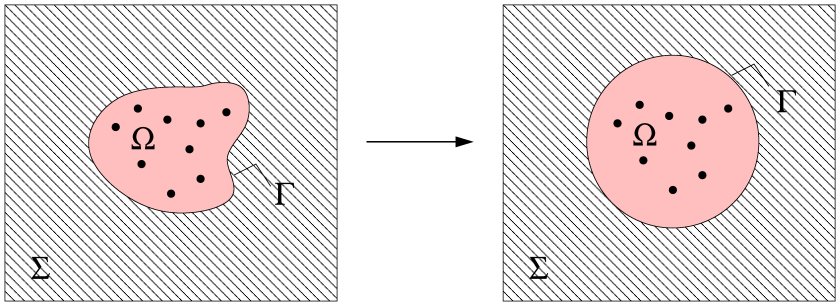


Fig. 6.6: The surface of the molecule is expanded to a sphere.

In order to justify this assumption, we analyze the solutions of an arbitrary charge distribution immersed in a sphere. We therefore reduce the complex shape of the molecule to a sphere and keep the charge distribution as illustrated in Fig. 6.6, i.e., completely contained by the sphere. The restriction to a spherical shape should not be a serious one for common globular proteins, because, even if the latter deviate somewhat from the spherical shape, electrostatic interactions depend primarily on the distance between the charges and how far within the molecule these charges are placed. The spherical symmetry allows to easily express the equations in spherical coordinates and to develop the solution of the differential system in a product ansatz of the form $U(r)\,Y(\theta,\phi)$. Further, as the principle of superposition holds, we only have to calculate the solution for a *single*, *arbitrarily located* point charge inside the sphere. From this, we can derive the solution for any point charge distribution.

In the spherical case, we can easily analyze the potential values for an arbitrary, fixed charge distribution and we will finally see, that the modifications of the potential values outside the molecule due to the different dielectric response inside the molecule are small.

We choose the coordinate system in a way that the center of the sphere lies in the origin. We further assume the point charge q to be placed at $\boldsymbol{r}_q$. Before we state the electrostatic potential of the point charge immersed inside the sphere, we have to introduce some notation:

Definition 6.5 Bessel functions:
We define i_n, the *modified spherical Bessel function of the first kind*

$$i_n(z) := \sqrt{\frac{\pi}{2z}} I_{n+\frac{1}{2}}(z), \quad z \in \mathbb{C}\,,$$

which is related to the *modified Bessel function of the first kind*

$$I_n(z) = \frac{1}{2\pi\imath} \oint e^{(\frac{z}{2})(t+\frac{1}{t})} t^{-n-1} dt, \quad z \in \mathbb{C}\,.$$

Analogously, the *modified spherical Bessel function* of the second kind

$$k_n(z) := \sqrt{\frac{\pi}{2z}} K_{n+\frac{1}{2}}(z), \quad z \in \mathbb{C}\,,$$

which is related to the *modified Bessel function of the second kind*

$$K_n(z) = \frac{1}{2\pi} \frac{I_{-n}(z) - I_n(z)}{\sin(n\pi)}, \quad z \in \mathbb{C}\,.$$

The modified spherical Bessel functions, $i_n(z)$ and $k_n(z)$, build up the basis of the solutions f of the differential equation

$$[r^2\partial_r^2 + 2r\partial_r - (r^2 + n(n+1))]f(r) = 0\,, \quad r \in \mathbb{R}\,,$$

which is related to the radial part of the Laplace operator in spherical coordinates [4, 19, 27, 76].

With Definition 6.5 we can set up the solutions for the two scalar nonlocal models applied to a spherical system:

$$4\pi\phi_\Omega(\boldsymbol{r}) = \sum_{n\geq 0}(\frac{1}{\varepsilon_\Omega}\frac{{r_q}^n}{r^{n+1}} + A_n r^n)P_n(\cos\alpha), \qquad \boldsymbol{r}\in\Omega \tag{6.12}$$

$$4\pi\psi_\Sigma(\boldsymbol{r}) = \sum_{n\geq 0} B_n \frac{1}{r^{n+1}}P_n(\cos\alpha), \qquad \boldsymbol{r}\in\Sigma \tag{6.13}$$

$$4\pi\phi_\Sigma(\boldsymbol{r}) = \frac{1}{\varepsilon_\Sigma}\psi_\Sigma(\boldsymbol{r}) + \sum_{n\geq 0} F_n\, k_n(r/\lambda)\, P_n(\cos\alpha), \qquad \boldsymbol{r}\in\Sigma\,, \tag{6.14}$$

where

$$\cos\alpha = \frac{\boldsymbol{r}\,\boldsymbol{r}_q}{r\,r_q}, \quad r = |\boldsymbol{r}|, r_q = |\boldsymbol{r}_q| \tag{6.15}$$

and $P_n(x)$ is the n^{th} Legendre polynomial [76]. All nonlocal models have the same functional structure, but differ in the amplitudes. In particular, we have for the DSM

$$\begin{aligned}
z &:= a/\lambda \\
B_0 &= 1 \\
B_{n\neq 0} &= \frac{{r_q}^n}{\varepsilon_\Omega}(2n+1)/\left(-(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma})\left(\frac{n(n+1)\,k_n(z)}{nk_n(z)-zk_{n+1}}\right)+\frac{1}{\varepsilon_\Omega}(n+1)+\frac{1}{\varepsilon_\Sigma}n\right) \\
A_0 &= \frac{1}{a}\left((\frac{1}{\varepsilon_\Sigma}-\frac{1}{\varepsilon_\Omega})+(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma})\frac{1}{1+z}\right) \\
A_{n\neq 0} &= \frac{1}{\varepsilon_\Omega}\frac{n+1}{n}({r_q}^n - B_n)/a^{2n+1} \\
F_0 &= \frac{2}{\lambda\pi}(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma})\frac{e^z}{1+z} \\
F_{n\neq 0} &= -B_n(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma})\frac{(n+1)a^{-(n+1)}}{nk_n(z)-zk_{n+1}(z)}\,,
\end{aligned}$$

where a defines the radius of the sphere. For the derivation we used the identity

$$\partial_x k_n(\kappa x) = \frac{n}{x}k_n(\kappa x) - \kappa k_{n+1}(\kappa x), \quad n\geq 0\,.$$

In Section 3.2.1 we introduced the correlation length λ, which defines the region of measurable hydrogen correlations. In general, the correlation length is comparable with the molecule's extension and this is why we consider $z = \frac{a}{\lambda}$ to be small. Then, the Bessel functions can be approximated [4]

$$k_n(z) \approx \frac{\pi}{2}\frac{(2n-1)!!}{z^{n+1}}, \qquad n>0,\, z<\sqrt{n+\frac{3}{2}}\,.$$

Using this approximation yields for the coefficients B_n and F_n,

$$\begin{aligned}
B_0 &= 1. \\
B_{n\neq 0} &= \underbrace{\frac{2n+1}{(\frac{\varepsilon_\Omega}{\varepsilon_\infty}+1)}}_{\approx 2}\frac{1}{n+1}{r_q}^n \approx {r_q}^n
\end{aligned} \tag{6.16}$$

and

$$\begin{aligned} F_0 &= \frac{2}{\lambda\pi}(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma})\frac{e^z}{1+z} \approx \frac{2}{\lambda\pi}(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma}) \\ F_{n\neq 0} &= \frac{2}{\lambda\pi}(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma})(2n+1)\underbrace{\left\{\frac{{r_q}^n}{\lambda^n}\frac{1}{(2n+1)!!}\right\}}_{i_n(r_q/\lambda)} . \end{aligned}$$

In Eq. (6.16) we claim that $\frac{\varepsilon_\Omega}{\varepsilon_\infty} \approx 1$, i.e., that ε_Ω and ε_∞ are of the same order of magnitude. This is a realistic assumption, as the macroscopic dielectric constant commonly used for proteins, ε_Ω ranges between $2\varepsilon_0$ and $10\varepsilon_0$ and $\varepsilon_\infty = 1.8\varepsilon_0$. The previous equations can then be combined to

$$\begin{aligned} B_n &= {r_q}^n && n \geq 0 \\ F_n &= \frac{2}{\lambda\pi}(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma})(2n+1)\, i_n(r_q/\lambda) && n \geq 0\,. \end{aligned}$$

Inserting these identities into Eqs. (6.13) and (6.14) yields the following approximation of the non-local potentials in Σ

$$\begin{aligned} 4\pi\psi_\Sigma(\boldsymbol{r}) &\approx \sum_{n\geq 0}\frac{{r_q}^n}{r^{n+1}}P_n(\cos\alpha)\,, && \boldsymbol{r}\in\Sigma \\ 4\pi\phi_\Sigma(\boldsymbol{r}) &\approx \frac{1}{\varepsilon_\Sigma}\psi_\Sigma + (\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma})\sum_{n\geq 0}\frac{2}{\lambda\pi}(2\,n+1)\, i_n(r_q/\lambda)\, k_n(r/\lambda)P_n(\cos\alpha)\,, && \boldsymbol{r}\in\Sigma\,. \end{aligned}$$

This result can be further simplified when comparing it to the expansion of the fundamental solutions of the Laplace $\triangle$, and the Yukawa operator $\mathcal{L}_{\frac{1}{\lambda}}$, in spherical harmonics [19, 52]:

$$\begin{aligned} \frac{1}{4\pi}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_q|} &= \sum_{n\geq 0}\frac{{r_q}^n}{r^{n+1}}P_n(\cos\alpha) \\ \frac{1}{4\pi}\frac{e^{-\frac{1}{\lambda}|\boldsymbol{r}-\boldsymbol{r}_q|}}{|\boldsymbol{r}-\boldsymbol{r}_q|} &= \sum_{n\geq 0}\frac{2}{\lambda\pi}(2\,n+1)\, i_n(r_q/\lambda)\, k_n(r/\lambda)P_n(\cos\alpha) \end{aligned}$$

Hence, we have

$$4\pi\psi_\Sigma(\boldsymbol{r}) \approx \frac{1}{|\boldsymbol{r}-\boldsymbol{r}_q|}\,, \qquad \boldsymbol{r}\in\Sigma \tag{6.17a}$$

$$4\pi\phi_\Sigma(\boldsymbol{r}) \approx \frac{1}{\varepsilon_\Sigma}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_q|} + (\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma})\frac{e^{-\frac{1}{\lambda}|\boldsymbol{r}-\boldsymbol{r}_q|}}{|\boldsymbol{r}-\boldsymbol{r}_q|}\,, \qquad \boldsymbol{r}\in\Sigma\,. \tag{6.17b}$$

Eqs. (6.17) are already known from the point charge solution given in Theorem 3.3.3. This proves the assumption made at the beginning of this section: the charge distribution directly immersed in water as illustrated in Fig. 6.5 yields a reasonable approximation for the boundary values of the electrostatic potentials. We summarize a few notes in the following remark.

Important remark 6.1

- In the presented derivation we considered a spherical surface. Whether this approximation is fulfilled or not depends on the molecule's composition and on the solvent. At least when assuming (polar) water and a mixture of nonpolar and polar side chains the molecules tend to cluster in order to

minimize the contact of the nonpolar side chains with the polar solvent, so that a spherical surface is macroscopically reasonable.

- The presented approximation does not exploit the limit $r \gg a$, which would be an additional constraint, when we calculate the boundary values for the finite difference box. We used that the correlation length is comparable or even larger than the extension of the molecule and that ε_Ω has the same order of magnitude as ε_∞. With these assumptions, we derived an analytical expression for the potentials *everywhere outside* the sphere. When approaching the surface of the molecule, our assumption to have a spherical molecule will certainly break down. However, as the approximated fields, Eqs. (6.17a-b) are the fundamental solutions of the nonlocal system *without* any dielectric boundary (see Remark 3.3 on p. 46), we suggest that the surface is in fact of minor importance in Σ.

- The presented field approximations given in Eqs. (6.17a-b) were derived explicitly for the DSM. However, a similar derivation can be made for the NSM as well as for the vector models. This can already be seen for the analytical expressions in the case of the Born sphere, where each of the nonlocal models results in Eqs. (6.17) for small $\frac{a}{\lambda}$ (Section 5.3). Thus, we expect all the nonlocal Lorentzian models to be similar in Σ, whereas in Ω they will differ from each other.

- An analogous derivation for the local model yields [19, 30]:

$$\begin{aligned} 4\pi\phi_\Omega(\boldsymbol{r}) &= \sum_{n\geq 0}\Big(\frac{1}{\varepsilon_\Omega}\frac{{r_q}^n}{r^{n+1}} + A_n r^n\Big)P_n(\cos\alpha) \\ 4\pi\phi_\Sigma(\boldsymbol{r}) &= \sum_{n\geq 0} B_n \frac{1}{r^{n+1}} P_n(\cos\alpha) \\ A_n &= \frac{{r_q}^n}{\varepsilon_\Omega}\frac{(\varepsilon_\Omega-\varepsilon_\Sigma)(n+1)}{(\varepsilon_\Omega n+\varepsilon_\Sigma(n+1))a^{2n+1}} \\ B_n &= \frac{{r_q}^n}{\varepsilon_\Sigma}\underbrace{\frac{2n+1}{(\frac{\varepsilon_\Omega}{\varepsilon_\Sigma}+1)\,n+1}}_{\approx 1} \approx \frac{{r_q}^n}{\varepsilon_\Sigma}\underbrace{\frac{2n+1}{n+1}}_{f(n)} \end{aligned} \tag{6.18}$$

  In order to result in the boundary approximations given in Eq. (6.17) we used $\frac{a}{\lambda} \ll 1$. The other limit, $\frac{a}{\lambda} \gg 1$, i.e., $\lambda \to 0$, yields the amplitudes B_n given in Eq. (6.18), which means that the local model is regained. As can be seen in Eq. (6.18), the amplitudes B_n of the local model still depend on n, because ε_Ω and ε_Σ are *not* of the same order of magnitude.

The derivation and the concluding remarks strongly support the choice of the following approximation for the boundary values of a finite difference calculation for the nonlocal Lorentzian models.

Definition 6.6 The approximated boundary condition:
The *approximated boundary condition* assumes the charge distribution (6.1) to be directly immersed in water, $\Sigma = \mathbb{R}^3$ and $\Omega = \emptyset$, i.e., skipping the boundaries and therefore neglecting the coupling of the two different dielectric media. The nonlocal electrostatic potentials of this setting read as follows

$$\psi_\Sigma(\boldsymbol{r}) = \frac{1}{4\pi}\int_{\mathbb{R}^3}\frac{\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\,d\boldsymbol{r}' \tag{6.19a}$$

$$\phi_\Sigma(\boldsymbol{r}) = \frac{1}{\varepsilon_\Sigma}\psi_\Sigma(\boldsymbol{r}) + \frac{1}{4\pi}\left(\frac{1}{\varepsilon_\infty}-\frac{1}{\varepsilon_\Sigma}\right)\int_{\mathbb{R}^3}\rho(\boldsymbol{r}')\frac{e^{-\frac{1}{\lambda}|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\,d\boldsymbol{r}' \tag{6.19b}$$

$$f_\Sigma(\boldsymbol{r}) = \frac{1}{\varepsilon_\Sigma-\varepsilon_\infty}\left(\psi_\Sigma(\boldsymbol{r})-\varepsilon_\infty\phi_\Sigma(\boldsymbol{r})\right). \tag{6.19c}$$

In order to *efficiently* generate boundary values, we simplify the charge distribution ρ by calculating an effective dipole:

Definition 6.7 The effective dipole boundary condition:
Assuming the arbitrary charge distribution ρ given in Eq. (6.1) on p. 85. For index sets $I^+ := \{i \,|\, 1 \leq i \leq N\,,\ q_i > 0\}$ and $I^- := \{i \,|\, 1 \leq i \leq N\,,\ q_i \leq 0\}$ we define

$$Q^{\pm} := \sum_{i \in I^{\pm}} q_i \quad , \quad \boldsymbol{r}^{\pm} := \frac{1}{Q^{\pm}} \sum_{i \in I^{\pm}} \boldsymbol{r}_i \, q_i \,,$$

and the effective (dipole) charge distribution

$$\rho_{\text{eff}} := Q^+ \, \delta(\boldsymbol{r} - \boldsymbol{r}^+) + Q^- \, \delta(\boldsymbol{r} - \boldsymbol{r}^-) \,.$$

Replacing the molecule's charge density by ρ_{eff} in Eqs. (6.19a-c), we gain a fast boundary value approximation denoted *effective dipole boundary condition*. If one of the effective charges is zero, only the other is taken into account and we obtain an effective monopole.

6.3 Surface generation

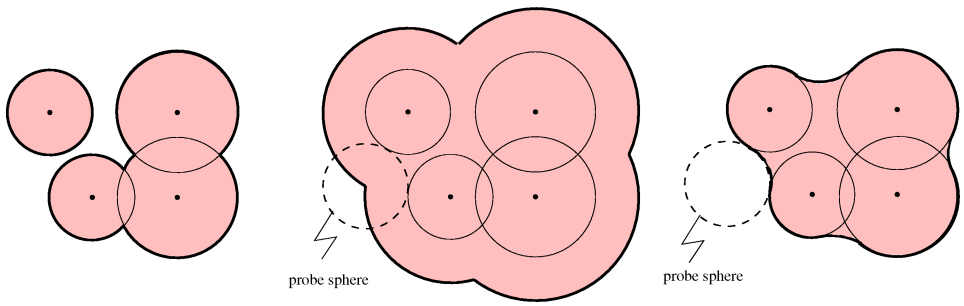
In Section 6.2 we developed an EJIIM solver and an appropriate boundary value approximation. In order to start an EJIIM calculation, it remains to provide the surface information of the molecule in the 3D grid. Besides the inside/outside information of every grid point, this includes the knowledge of the intersection points between the Cartesian grid and the molecular surface Γ together with the normal and the tangential vectors at these intersection points on Γ. In the literature, very accurate and efficient methods to calculate the inside/outside information are discussed [101]. However, these algorithms provide neither the exact intersection points nor information on the surface normal direction. In this section we derive an algorithm to generate a discrete surface description of a biomolecule in an equidistant 3D Cartesian grid.

After starting with basic notations in Section 6.3.1, we set up the algorithm to calculate the VdWS and the SAS. The algorithm is described in Section 6.3.2. Based on these results, we calculate the molecular surface, i.e., the SES of the molecule, in Section 6.3.3.

In the first instance, the algorithm was developed to generate the input information for the EJIIM. In addition, we can easily use the discrete surface description to calculate a triangulation for the BEM. This extension is presented in Section 6.3.5.

6.3.1 Definitions

In Sections 3.1.2 we introduced the basic notations to describe the surface of a molecule.



(a) VdWS (b) SAS (c) SES

The three different types of surfaces used in biomolecular research, the VdWS, the SAS, and the SES have been illustrated in Fig. 3.3, which is once again shown above. They were described on p. 26. Now, we specify these definitions in a theoretical way. This is the basis for efficiently defining the intersections of the molecule with a 3D grid.

Definition 6.8 Surface definition
Assume a molecule $\mathcal{M}$ to consist of a set of N atoms. Each is represented as a Born sphere S with given center $\boldsymbol{r}_i$ and radius R_i, $i = 1, \ldots, N$. We define the domain $\mathcal{M}$ as

$$\mathcal{M} = \bigcup_{i=1}^{N} S(\boldsymbol{r}_i, R_i)\,. \tag{6.20}$$

We introduce ∂V as the boundary of a domain $V \subset \mathbb{R}^3$ and we further denote the radius of a solvent molecule as r^{probe}. With these notations, we define

(a) the *Van der Waals Surface* (VdWS):

$$\mathrm{VdWS}(\mathcal{M}) = \partial\left(\bigcup_{i=1}^{N} S(\boldsymbol{r}_i, R_i)\right) = \partial\mathcal{M} \tag{6.21}$$

(b) the *Solvent Accessible Surface* (SAS):

$$\mathrm{SAS}(\mathcal{M}) = \partial\left(\bigcup_{i=1}^{N} S(\boldsymbol{r}_i, R_i + r^{probe})\right) \tag{6.22}$$

(c) the *Solvent Excluded Surface* (SES):

$$\mathrm{SES}(\mathcal{M}) = \partial\left(\bigcup_{i=1}^{N} S(\boldsymbol{r}_i, R_i + r^{probe}) \Big\backslash \bigcup \left(S(\boldsymbol{r}, r^{probe})\,, \boldsymbol{r} \in \mathrm{SAS}(\mathcal{M})\right)\right) \tag{6.23}$$

When analyzing these three different surface definitions, it is remarkable that each of them is based on a *merging* or an *exclusion of spheres*. This is in fact the basis of the algorithm: we will implement a merging procedure to consecutively build the surface of the molecule out of spheres.

Before we give the details on the algorithm, we have to agree on the following definitions.

Along the lines of Definition 6.1 on p. 89 we introduce a regular grid g with spacing h and size $(n_x \times n_y \times n_z)$. We further keep the standard notation $g_{ijk} := g(x_i, y_j, z_k)$ for a grid point (x_i, y_j, z_k), $i \in \{1, \ldots, n_x\}$, $j \in \{1, \ldots, n_y\}$, $k \in \{1, \ldots, n_z\}$.

Points on the edges of the grid which belong to the surface ∂V of a volume V are called *intersections* of V. To avoid pathological cases and to allow for certain precalculations, intersection points are restricted to lie on discrete, equidistant distances on the edges of the grid, i.e., we discretize the set of intersection points. This is shown in Fig. 6.7 on the left, where the small black points on the grid lines indicate all possible positions of intersection points. Typically, we allow for ten possible intersection positions along the distance h. Additionally, only a single intersection point is considered on the edge between two adjacent grid points, in a way that one of these grid points lies in V, whereas the other lies in the complement $\overline{V}$. If the exact intersection does not fulfill this

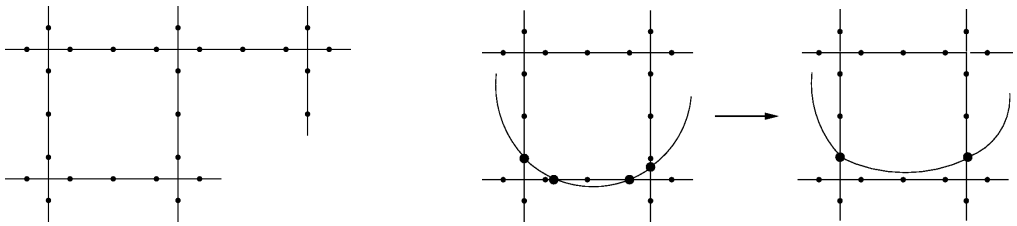


Fig. 6.7: Restrictions: (left) discretization of the intersection point position; (right) local deformation of the surface to guarantee a single intersection point on the edges of the grid.

restriction, we deform the surface as shown on the right of Fig. 6.7. Such a surface correction is on a length scale below the grid resolution and thus, does not spoil the numerical results.

In order to efficiently describe the grid-based surface we now introduce special sets of grid points:

Definition 6.9 Point sets:
We consider a volume $V \subset \mathbb{R}^3$ with surface ∂V and complement $\overline{V}$. V completely lies in a regular grid g with grid spacing h. Then, all grid points g_{ijk} of the regular grid are assigned to one of the following grid point sets

$$\begin{aligned} \gamma_{V,I} &:= \{g_{ijk} \mid 0 < |g_{ijk} - \partial V| < h \text{ and } g_{ijk} \in V\} \\ \gamma_{\overline{V},I} &:= \left\{g_{ijk} \mid 0 < |g_{ijk} - \partial V| < h \text{ and } g_{ijk} \in \overline{V}\right\} \\ V_I &:= \{g_{ijk} \mid g_{ijk} \in V \text{ and } g_{ijk} \notin \gamma_{V,I}\} \\ \overline{V}_I &:= \left\{g_{ijk} \mid g_{ijk} \in \overline{V} \text{ and } g_{ijk} \notin \gamma_{\overline{V},I}\right\} , \end{aligned}$$

where $|g_{ijk} - \partial V|$ denotes the distance of the grid point g_{ijk} to the next point lying on an edge of the grid and being part of ∂V.

Tab. 6.1 introduces predefined notations for the volume V in the case of the molecule and in the case of a sphere. The molecule is described by the point sets $\{\Omega_I, \gamma_{\Omega,I}, \Sigma_I, \gamma_{\Sigma,I}\}$, whereas the sphere is defined by $\{\mathrm{S}_I, \gamma_{\mathrm{S},I}, \overline{\mathrm{S}}_I, \gamma_{\overline{\mathrm{S}},I}\}$.

Tab. 6.1: Notations.

	V	$\overline{V}$	∂V
molecule	Ω	Σ	Γ
sphere	S	$\overline{\mathrm{S}}$	∂S

By means of a cross section through a single atom sphere S, Fig. 6.8 illustrates the meaning of the different grid point sets of Definition 6.9. Grid points which lie outside the sphere are drawn as white circles. They can either contribute to $\gamma_{\overline{\mathrm{S}},I}$ or to $\overline{\mathrm{S}}_I$. In contrast, the inner grid points are colored in gray. All inner points either belong to $\gamma_{\mathrm{S},I}$ or to S_I. For the grid points shown in the inset, we listed the membership in detail in the caption of Fig. 6.8.

Edges connecting a grid point of $\gamma_{V,I}$ with a grid point of $\gamma_{\overline{V},I}$ are additionally drawn as solid lines in the inset of Fig. 6.8. Since all intersection points lie on these edges, they are denoted *non-trivial* in contrast to the *trivial* edges, which connect grid points lying both either in V or in $\overline{V}$.

In order to uniquely characterize the intersections lying on the non-trivial edges, we assign to each grid point in $\gamma_{V,I}$ and in $\gamma_{\overline{V},I}$ the so called *edge label*, which is specified in Definition 6.10.

Definition 6.10 Edge label:
The edge label $e_{d,ijk}$ refers to an edge originating at grid point g_{ijk} and connecting the next grid point in direction $\boldsymbol{e}_d \in \{\pm\boldsymbol{e}_x, \pm\boldsymbol{e}_y, \pm\boldsymbol{e}_z\}$.

For non-trivial edges, the edge label is comprised of the relative distance d where the intersection point is found and a set $\mathcal{S}_{spheres}$ of all sphere indices of spheres that contribute to the respective intersection point.

We further define

$$N(\mathrm{e}_{ijk}^{\text{non-trivial}}) := \text{number of non-trivial edges the grid point } g_{ijk} \text{ participates in}\,.$$

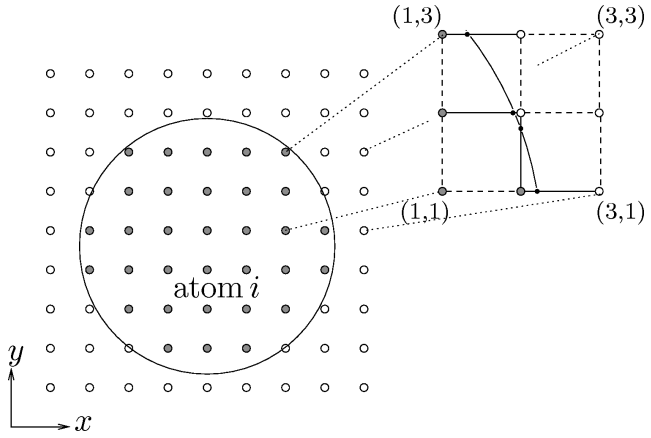


Fig. 6.8: Grid based representation of an atom sphere: dark-gray grid points lie in S, white ones in $\overline{\mathrm{S}}$. Inset: grid points divide into the following sets: $\mathrm{S}_I = \{(1,1)\}$, $\gamma_{\mathrm{S},I} = \{(1,2),(1,3),(2,1)\}$, $\gamma_{\overline{\mathrm{S}},I} = \{(2,2),(2,3),(3,1)\}$, $\overline{\mathrm{S}}_I = \{(3,2),(3,3)\}$.

Tab. 6.2: Complete edge label of grid point $(2,2)$ in Fig. 6.8.

edge	(2,2)-(3,2)	(2,2)-(1,2)	(2,2)-(1,3)	(2,2)-(2,1)
$\boldsymbol{e}_d$	$\boldsymbol{e}_x$	$-\boldsymbol{e}_x$	$\boldsymbol{e}_y$	$-\boldsymbol{e}_y$
d	/	0.1	/	0.2
$\mathcal{S}_{spheres}$	/	$\{i\}$	/	$\{i\}$
type	trivial	non-trivial	trivial	non-trivial

In order to give an example for the complete edge label of a grid point, we consider in Tab. 6.2 all edges of the grid point $(2,2)$ of Fig. 6.8. For instance, the edge, which connects $(2,2)$ with $(2,1)$, is non-trivial, because $(2,2) \in \gamma_{\overline{V},I}$ and $(2,1) \in \gamma_{V,I}$. With the edge label given in Tab. 6.2 we identify the position (x,y) of the intermediate intersection point in Cartesian coordinates as

$$(x,y) = (x_2,y_2) + \underbrace{0.2\,(0,\text{-}1)}_{d\boldsymbol{e}_d}\, h\ .$$

Furthermore, we see in Fig. 6.8 that the grid point $(2,2)$ participates in two non-trivial edges whereas, for example, the grid point $(1,2)$ belongs to only one non-trivial edge. Thus, we have $N(\mathrm{e}_{12}^{\text{non-trivial}}) = 1$ for $(1,2)$ and $N(\mathrm{e}_{22}^{\text{non-trivial}}) = 2$ for $(2,2)$.

Since Fig. 6.8 illustrates a single atom sphere, whose atom information (e.g., radius and center) can be retrieved from its atom index i, the set $\mathcal{S}_{spheres}$ of all non-trivial edges consist of i, only. In contrast, in Fig. 6.9 the intersection point on the lined, non-trivial edge belongs to two different atoms, namely i and j and this yields the set $\mathcal{S}_{spheres} = \{i,j\}$ for the lined edge.

The *trivial* edges do not retain intersection or atom connectivity information and thus we do not store them explicitly. With the knowledge of all non-trivial edges, the trivial ones are easily detectable, because both sets exclude each other. In the inset of Fig. 6.8, all trivial edges are dashed such as the edge that connects the grid points (2,2) and (3,2).

Based on the above definitions and constraints, the complete intersection information is contained in both $\gamma_{V,I}$ and $\gamma_{\overline{V},I}$. This allows for a fast access to intersection information and atom connectivity.

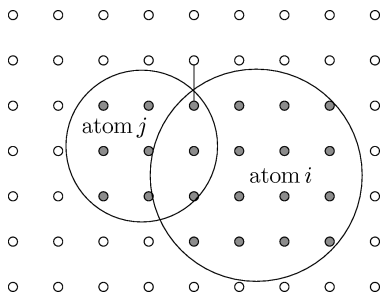


Fig. 6.9: The set $\mathcal{S}_{spheres}$ of the lined non-trivial edge consists of the atoms with indices i and j.

We specify the set of intersection points, Γ_I in Definition 6.11.

Definition 6.11 Discrete set of intersection points

$$\begin{aligned}\Gamma_I &:= \left\{(x,y,z) = (g_{ijk} + d\,\boldsymbol{e}_d) \in \mathbb{R}^3 \mid g_{ijk} \in \gamma_{V,I} \text{ and } e_{d;ijk} \text{ is non-trivial}\right\} \\ &:= \left\{(x,y,z) = (g_{ijk} + d\,\boldsymbol{e}_d) \in \mathbb{R}^3 \mid g_{ijk} \in \gamma_{\overline{V},I} \text{ and } e_{d;ijk} \text{ is non-trivial}\right\}\end{aligned}$$

By construction, the grid-based representation of an arbitrarily shaped volume V can be uniquely described by $\{V_I, \gamma_{V,I}\}$ *or* $\{\overline{V}_I, \gamma_{\overline{V},I}\}$, where $\gamma_{V,I}$ or $\gamma_{\overline{V},I}$ refers to the non-trivial edges with respect to V or $\overline{V}$. For the grid-based SAS generation, we decided to describe the surfaces of the atom spheres and of the molecule with the inside information, i.e., with $\{V_I, \gamma_{V,I}\}$[3].

As a special case, we introduce $\{\mathrm{S}_I^k, \gamma_{\mathrm{S},I}^k\}$ for the grid-based representation of a sphere with index k and $\{\Omega_I^{(k:j)}, \gamma_{\Omega,I}^{(k:j)}\}$ as the grid based representation of an union of the k^{th} to the j^{th} sphere of the molecule.

Finally, we define the *atom sphere*, which has already been implicitly used in the last illustrations:

Definition 6.12 Atom sphere:
The *atom sphere* is defined by the point sets $\{\mathrm{S}_I, \gamma_{\mathrm{S},I}\}$ or $\{\overline{\mathrm{S}}_I, \gamma_{\overline{\mathrm{S}},I}\}$. Additionally, it holds, that the inside information of the sphere equals the inside information of the molecule, also denoted *molecule inside information*. Thus, an accumulation of atom spheres belongs to Ω.

In all previous and also in the coming illustrations, dark-gray (white) points indicate grid points, which lie inside (outside) the molecule or inside (outside) an atom sphere. At the moment, the definition of the atom sphere seems self-evident and therefore superfluous, however, when we construct the SES in Section 6.3.3 we require the *probe sphere*, which has an inside/outside information contrary to the molecule and to the atom sphere, i.e., the inner of a probe sphere bears the outside information of the molecule, also denoted *molecule outside information* (see Definition 6.13).

6.3.2 Construction of the grid-based SAS

In this section, we explain the algorithm to generate the grid-based SAS information of a molecule. As already motivated above, the construction of the grid-based SAS is done by the consecutive merging of atom spheres as defined in Eq. (6.22). We assume the l^{th} step of the merging procedure, where the sphere information of the l^{th} atom sphere is merged with the $(l-1)$ already merged atom

[3]Choosing the outside information, $\{\overline{V}_I, \gamma_{\overline{V},I}\}$ would change the $\oplus_{\mathrm{A}}$-operator (Tab. 6.3(b)) in a way that the operator minimizes the outer domain, i.e., to the $\oplus_{\mathrm{P}}$-operator (Tab. 6.4).

spheres of the molecule,

$$\underbrace{\Big(S(\boldsymbol{r}_l, R_l + r^{probe})\Big)}_{l^{th}\text{ atom sphere to be merged}} \cup \underbrace{\left(\bigcup_{i=1}^{l-1} S(\boldsymbol{r}_i, R_i + r^{probe})\right)}_{(l\text{–}1)\text{ already merged atom spheres}} . \tag{6.24}$$

With the definitions made in Section 6.3.1, the merging procedure computes the new set $\{\Omega_I^{(1:l)},\, \gamma_{\Omega,I}^{(1:l)}\}$ from given sets $\{\mathrm{S}_I^l,\, \gamma_{\mathrm{S},I}^l\}$ and $\{\Omega_I^{(1:l-1)},\, \gamma_{\Omega,I}^{(1:l-1)}\}$. This is described in pseudo-code in Algorithm 6.1.

Algorithm 6.1 Merging atom spheres

Require: $\{\mathrm{S}_I^1, \gamma_{\mathrm{S},I}^1\}$ or $\Big(\{\Omega_I^{(1:l-1)}, \gamma_{\Omega,I}^{(1:l-1)}\}$ and $\{\mathrm{S}_I^l, \gamma_{\mathrm{S},I}^l\}\Big)$.

Ensure: $\{\Omega_I^{(1:l)}, \gamma_{\Omega,I}^{(1:l)}\}$.

1: **if** $l = 1$ **then**
2: initialize: $\{\Omega_I^{(1:l)}, \gamma_{\Omega,I}^{(1:l)}\} = \{\mathrm{S}_I^1, \gamma_{\mathrm{S},I}^1\}$
3: **else**
4: initialize: $\{\Omega_I^{(1:l)}, \gamma_{\Omega,I}^{(1:l)}\} = \{\Omega_I^{(1:l-1)}, \gamma_{\Omega,I}^{(1:l-1)}\}$
5: **for all** $g_{ijk} \in \mathrm{S}_I^l$ **do**
6: insert g_{ijk} in $\Omega_I^{(1:l)}$
7: **if** $g_{ijk} \in \gamma_{\Omega,I}^{(1:l)}$ **then**
8: remove g_{ijk} from $\gamma_{\Omega,I}^{(1:l)}$
9: **for all** $g_{ijk} \in \gamma_{\mathrm{S},I}^l$ **do**
10: **for all** $\boldsymbol{e}_d \in \{\pm\boldsymbol{e}_x, \pm\boldsymbol{e}_y, \pm\boldsymbol{e}_z\}$ **do**
11: $\mathrm{e}_{d;ijk}^l \oplus_{\mathrm{A}} \mathrm{e}_{d;ijk}^{(1:l)} = \left\{ \begin{array}{lll} 1 & : \text{insert} & \mathrm{e}_{d;ijk}^l \\ 1^{max} & : \text{update} & \mathrm{e}_{d;ijk}^{(1:l)} \\ 0 & : \text{remove} & \mathrm{e}_{d;ijk}^{(1:l)} \end{array} \right\}$ in $\gamma_{\Omega,I}^{(1:l)}$
12: **if** $\Big(\Big(N(\mathrm{e}_{ijk}^{\text{non-trivial}}) == 0\Big)$ for $g_{ijk} \in \gamma_{\Omega,I}^{(1:l)}\Big)$ **then**
13: remove g_{ijk} from $\gamma_{\Omega,I}^{(1:l)}$
14: insert g_{ijk} in $\Omega_I^{(1:l)}$
15: **return** $\{\Omega_I^{(1:l)}, \gamma_{\Omega,I}^{(1:l)}\}$

Description of Algorithm 6.1

$\boxed{l = 1}$

When merging the first sphere, the molecule set is empty and the set $\{\mathrm{S}_I^1, \gamma_{\mathrm{S},I}^1\}$ contains all the information about the current (molecular) surface $\{\Omega_I^{(1:1)}, \gamma_{\Omega,I}^{(1:1)}\}$. Thus, we are done with

the initialization given in line 2 of Algorithm 6.1.

$\boxed{l > 1}$

In order to generate the new set $\{\Omega_I^{(1:l)}, \gamma_{\Omega,I}^{(1:l)}\}$ we proceed in three steps:

- we initialize the current molecule set $\{\Omega_I^{(1:l)}, \gamma_{\Omega,I}^{(1:l)}\}$ with the old set $\{\Omega_I^{(1:l-1)}, \gamma_{\Omega,I}^{(1:l-1)}\}$.
- we unify the inside information of the sphere, S_I^l with the one of the current molecule, $\Omega_I^{(1:l)}$ as done in line 6 of Algorithm 6.1.
  - the union of $\Omega_I^{(1:l)}$ and S_I^l implies the deletion of grid points $g_{ijk} \in \gamma_{\Omega,I}^{(1:l)}$ that now belong to $\Omega_I^{(1:l)}$, as the sets have to exclude each other. This is done in lines 7-8 of Algorithm 6.1.
- we merge $\gamma_{\mathrm{S},I}^l$ and $\gamma_{\Omega,I}^{(1:l)}$. In principle, we have to find all new non-trivial edges of the molecule's surface and we have to update old, non-trivial edges, which are covered by the surface of the l^{th} sphere. To this end, we iterate over all edges of grid points in $\gamma_{\mathrm{S},I}^l$ and proceed as follows:
  - we check whether the edge contributes to the molecular surface. If this is the case, we update the edge label in agreement with the constraints made in Section 6.3.1. This can mean that we insert the grid point g_{ijk} in $\gamma_{\Omega,I}^{(1:l)}$, or that we update its edge label in $\gamma_{\Omega,I}^{(1:l)}$, or even that we remove g_{ijk} from $\gamma_{\Omega,I}^{(1:l)}$. These steps are done in line 11 of Algorithm 6.1, where the $\oplus_A$-operator is defined by Tab. 6.3(b) and separately described in Section 6.3.2.1.
  - we remove the point $g_{ijk} \in \gamma_{\Omega,I}^{(1:l)}$, when g_{ijk} does not possess any non-trivial edges after the merging procedure. As in this case, g_{ijk} lies in Ω, the grid point is added to $\Omega_I^{(1:l)}$. This is realized in lines 13-14 of Algorithm 6.1.

Fig. 6.10 illustrates the molecule surface (lined) before and after the merging of three different atom spheres (dashed).

After the N^{th} atom sphere of the molecule has been processed and merged, the complete SAS information is stored in the sets $\{\Omega_I^{(1:N)}, \gamma_{\Omega,I}^{(1:N)}\}$, where $\gamma_{\Omega,I}^{(1:N)}$ contains the non-trivial edge label, i.e., the intersection and the atom connectivity information. The latter is required to calculate the normal vector on the surface at the intersection point: assume to have a set of L atoms that contribute to the intersection point $\gamma \in \Gamma_I^{(1:N)}$, then we define the outer normal vector $\boldsymbol{n}$ by

$$\boldsymbol{n} = \frac{1}{L} \sum_{i=1}^{L} \frac{\boldsymbol{r}_i - \boldsymbol{\gamma}}{|\boldsymbol{r}_i - \boldsymbol{\gamma}|}\,, \tag{6.25}$$

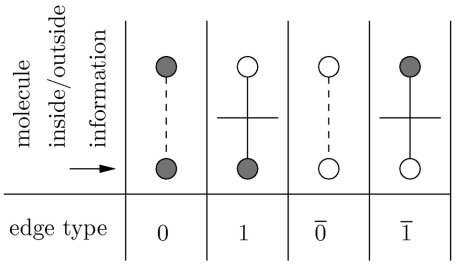
where $\boldsymbol{r}_i$ is the center of the i^{th} atom and $\boldsymbol{\gamma}$ is the position of the intersection point. The grid-based SAS is illustrated for a cross section through trypsin in Fig. 6.11(b).

Note that the construction of the grid-based VdWS is analogous to the SAS generation when r^{probe} is set to zero in Eq. (6.24).

6.3.2.1 Addition table: $\oplus_A$-operator

In order to describe the resulting edge configuration within the merging process of a sphere and a molecule edge, we assign the so called *edge type* to every edge. The edge type specifies the inside/outside information of the grid points of an edge. All possible inside/outside configurations are shown in Tab. 6.3(a):

(a) Molecule edge types



(b) Addition operator $\oplus_A$

$e^l \oplus_A e^{(1:l)}$		0	1	$\overline{0}$	$\overline{1}$
	0	0	0	0	0
	1	0	1^{max}	1	0

Tab. 6.3: (a) Edge types: a molecule edge type is allocated to every edge starting at the grid point which is denoted by the black arrow. It characterizes the inside/outside information of the edge: dark-gray grid points have molecule inside information, i.e., they lie in Ω (or in S), white ones have molecule outside information, i.e., they lie in Σ (or in $\overline{\mathrm{S}}$); (b) Addition operator $\oplus_A$: merging atom spheres with the $\oplus_\mathrm{A}$-operator; e^l denotes an edge of the atom sphere (rows) and $\mathrm{e}^{(1:l)}$ the corresponding edge of the molecule under construction (columns).

The edge type 0 denotes an edge with molecule inside information at both ends, whereas an edge with edge type $\overline{0}$ has molecule outside information at both ends. Edges with type 0 and $\overline{0}$ lie in Ω and Σ, or in the case of an atom sphere in S and $\overline{\mathrm{S}}$, respectively.

In contrast to these trivial edges with the same inside/outside information at both ends, one of the grid points of a non-trivial edge has inside the other has outside information. For the non-trivial edges, we introduce the edge type 1, when the reference grid point has inside information indicated by the black arrow in Tab. 6.3(a). In the mirrored case, the edge has type $\overline{1}$. Thus, by definition, edges with type 1 lie in $\gamma_{\Omega,I}$ or in the case of an atom sphere in $\gamma_{\mathrm{S},I}$, whereas edges with type $\overline{1}$ lie in $\gamma_{\Sigma,I}$ or, in the case of an atom sphere, in $\gamma_{\overline{\mathrm{S}},I}$.

Based on these edge types, we define the addition operator $\oplus_A$ in Tab. 6.3(b), which is used in the Algorithm 6.1 in line 11 to abbreviate the merging of an edge of an atom sphere, e^l with the corresponding edge of the molecule, $e^{1:l}$.

As we see in line 9 of Algorithm 6.1, the addition operation acts only on edges of grid points $g_{ijk} \in \gamma^l_{\mathrm{S},I}$ and this is why we consider only the edge types 0 and 1 in the rows of Tab. 6.3(b). In order to account for all edge types a sphere edge can face when merging with the corresponding molecule edge, we consider all four possible edge types in the columns.

In principle, the addition table is based on maximizing the surface with respect to the added sphere. Thus, the edge type 0, which is an edge with inside information, has the highest priority and every merging procedure where the edge type 0 is involved results in an edge with type 0 (e.g., $e_{+y;64}$ and $e_{+x;86}$ in Fig. 6.10). In contrast, the edge type $\overline{0}$, which denotes an edge with outside information, has the lowest priority and is always replaced in favor of the edge type of the added atom sphere (e.g., $e_{-y;33}$ in Fig. 6.10).

When an atom sphere edge of type 1 faces a molecule edge of type 1, the edge with the greater distance d to the intersection point defines the new surface, i.e., if necessary the old edge label is overwritten (cf. $e_{-x;63}$ and $e_{+y;77}$ in Fig. 6.10). This is implied by 1^{max} in Tab. 6.3(b). Otherwise, if the distances are equal, the molecule edge remains unchanged, but the sphere index l has to be added to the atom set $\mathcal{S}_{spheres}$ of this edge (e.g., $e_{+x;85}$ in Fig. 6.10).

Furthermore, the addition scheme always fulfills the restrictions that are made in Section 6.3.1. For instance, when an atom sphere edge of type 1 faces a molecule edge of type $\overline{1}$, the corresponding

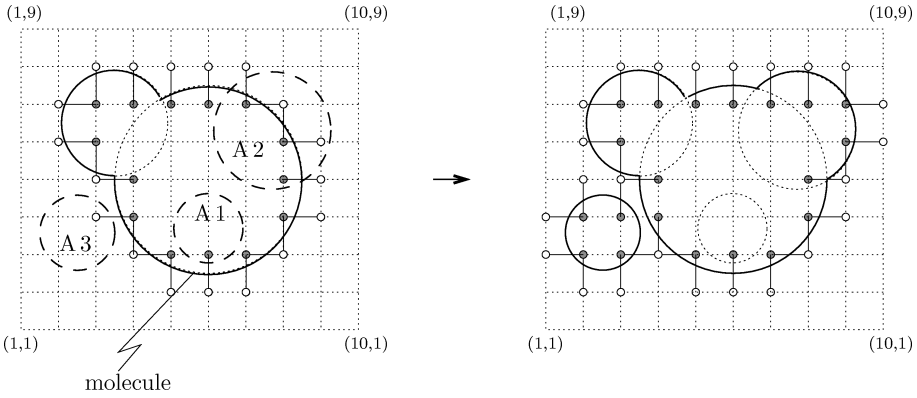


Fig. 6.10: Merging scenarios for the dashed atom spheres (1-3): A 1 lies inside the molecule, A 2 overlaps with the molecule's surface and A 3 lies within one grid spacing next to the molecule.

grid point $g_{ijk} \in \gamma_{\Omega,I}^{(1:l)}$ is skipped in either case (e.g., $e_{-x;44}$ in Fig. 6.10). After this procedure, the edge lies completely in Ω.

6.3.2.2 Acceleration of the algorithm

In order to accelerate the algorithm, the merging procedure is done for atoms lying on the surface, only. For the identification of all the atoms, which contribute to the VdWS or the SAS, we use the *reduced surface* implementation in the molecular software package BALL. The reduced surface is an analytical description of the molecular surface, and among other things it provides a list of the atoms of the molecule with a non-vanishing surface contribution [109]. Processing only the atoms of this list reduces the number of merging procedures considerably. For instance, from the total number of 3223 atoms the protein trypsin is composed of, only 1403 atoms are surface exposed. Fig. 6.11(a) exemplifies the result of the merging procedure for a cross section through trypsin.

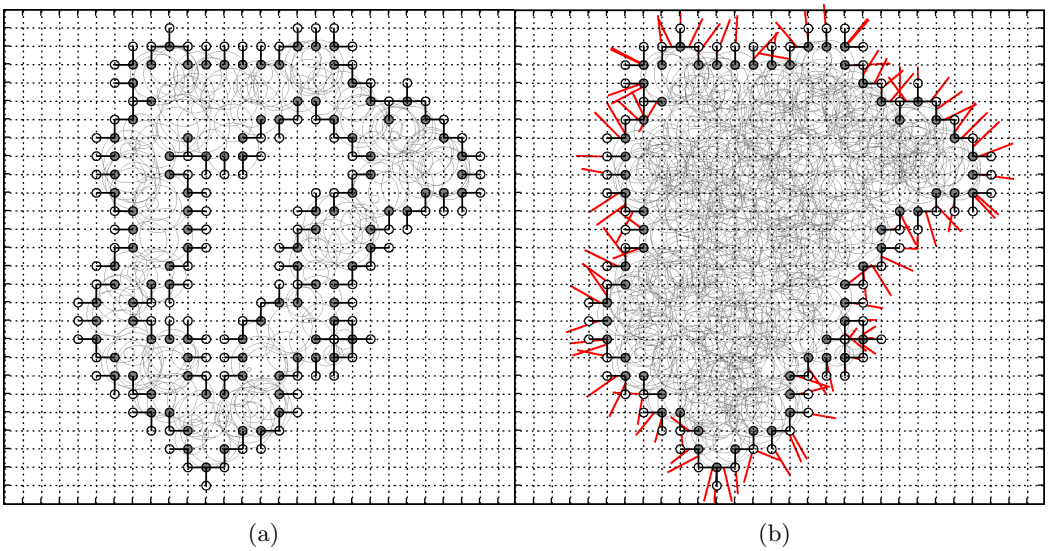


Fig. 6.11: (a) merging the surface exposed atoms; (b) grid-based representation of the SAS of trypsin after applying the flood-fill algorithm, red arrows denote the normal information.

Since the inner atoms have not been processed yet, there is a wrong inner band of points in $\gamma_{\Omega,I}^{1:N}$. In order to delete these points and to assign all inner points to the set Ω_I, we apply a standard, recursive inner flood-fill algorithm that starts at every buried atom position, i.e., at all atoms that do not contribute to the surface [54]. The inner flood-fill algorithm terminates, when reaching grid points that belong to the set $\gamma_{\Omega,I}^{1:N}$ [4].

6.3.3 Construction of the grid-based SES

First of all, we transfer the definition of the SES given in Eq. (6.23),

$$\mathrm{SES}(\mathcal{M}) = \partial\left(\bigcup_{i=1}^{N} S(\boldsymbol{r}_i, R_i + r^{probe}) \Big\backslash \bigcup\Big(S(\boldsymbol{r}, r^{probe}), \boldsymbol{r} \in \mathrm{SAS}(\mathcal{M})\Big)\right),$$

into a discretized context

$$\mathrm{SES} \;=\; \partial\left(\Omega_{I,\mathrm{SAS}} \Big\backslash \bigcup\left(S(\Gamma_{I,\mathrm{SAS}}, r^{probe})\right)\right), \tag{6.26}$$

where $\bigcup\left(S(\Gamma_{I,\mathrm{SAS}}, r^{probe})\right) := \bigcup\left(S(\boldsymbol{r}, r^{probe}), \boldsymbol{r} \in \Gamma_{I,\mathrm{SAS}}\right)$ describes the union of all *probe spheres* positioned at the discrete intersection points of the SAS. From this definition, it is directly clear that the SES generation consists of two algorithms. The first computes the inside information of the grid-based SAS of the molecule $\Omega_{I,\mathrm{SAS}}$ and the set of discrete SAS intersection points $\Gamma_{I,\mathrm{SAS}}$. The second algorithm imitates the rolling process in a discrete way by a consecutive merging of probe spheres which are located on the intersection points $\boldsymbol{r}_i \in \Gamma_{I,\mathrm{SAS}}$. The basic computation is depicted in Fig. 6.12, where we start with the set $\Sigma_{I,\mathrm{SAS}}$ and consecutively merge $\Sigma_{I,\mathrm{SAS}}$ with probe spheres centered at the points $\in \Gamma_{I,\mathrm{SAS}}$.

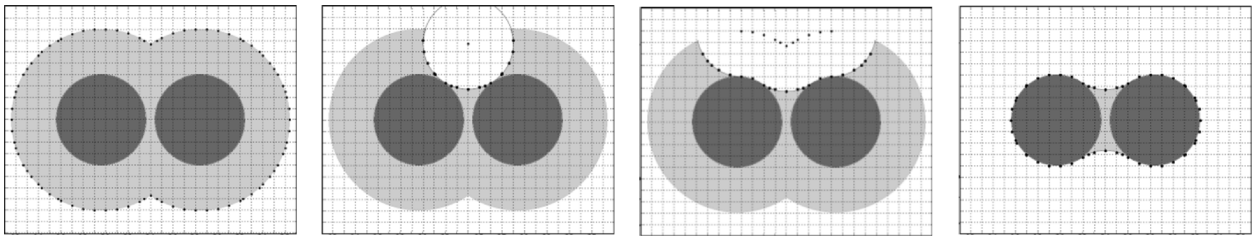


Fig. 6.12: left: VdW(dark gray), Ω_{SAS}(light gray), Σ_{SAS} (white), $\Gamma_{I,\mathrm{SAS}}$ (black points on SAS); middle: constructing SES by merging of probe sphere with $\Sigma_{I,\mathrm{SAS}}$: SAS intersection points that were already processed are additionally drawn as small black points; right: Σ_{SES} (white) and Ω_{SES} (the rest) after the complete merging procedure.

In order to adopt the complete merging procedure described in Section 6.3.2, we transform the $\backslash$-operator in Eq. (6.26) for the inner domain $\Omega_{I,\mathrm{SAS}}$ into the $\cup$-operator for the outer domain $\Sigma_{I,\mathrm{SAS}}$:

$$\mathrm{SES} \;=\; \partial^{inner}\left(\Sigma_{I,\mathrm{SAS}} \cup \left(\bigcup S(\Gamma_{I,\mathrm{SAS}}, r^{probe})\right)\right). \tag{6.27}$$

To demonstrate this equality we use the algebra of sets, in particular De Morgan's law, which is

[4] In order to assign "holes" inside the molecule to Ω, we have to start an outer flood-fill algorithm instead, meaning that we start at the border of the 3D grid. The outer flood-fill algorithm terminates, when grid points are reached, which belong to $\gamma_{\Omega,I}^{1:N}$. In fact, this algorithm is used for electrostatic problems, because the solvent does not act as a dielectric continuum inside inner holes and does not contribute to the water network.

illustrated in Fig. 6.13 for two set A and B, where in our case it holds

$$A = \Omega_{I,\mathrm{SAS}}, \quad \overline{A} = \Sigma_{I,\mathrm{SAS}}, \quad B = \bigcup S(\Gamma_{I,\mathrm{SAS}}, r^{probe}) \,.$$

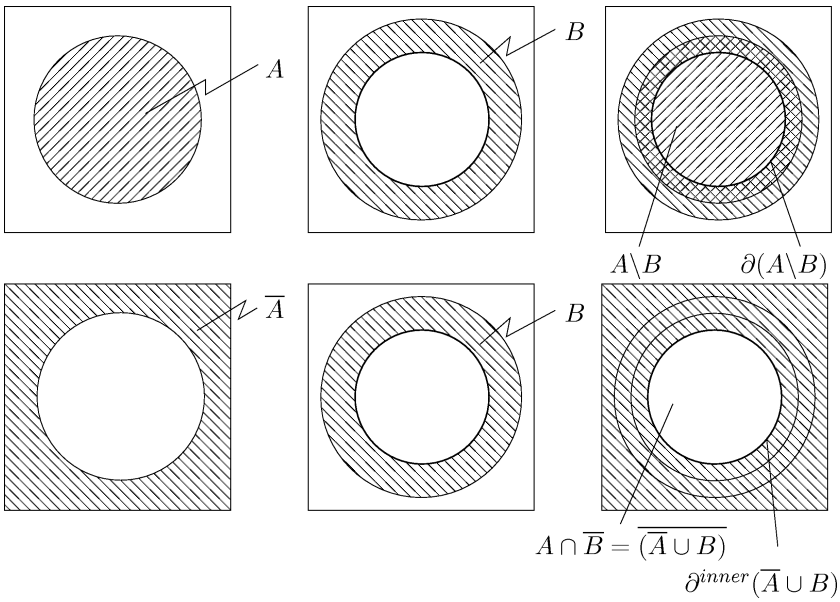
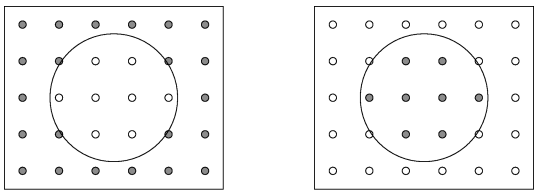


Fig. 6.13: Set algebra to reformulate the SES generation into a pure merging procedure.

In Definition 6.13 we formally introduce the probe sphere in analogy to the atom sphere. An illustration of the probe sphere is given in Fig. 6.14.

Definition 6.13 Probe sphere:
The *probe sphere* is defined by the point sets $\{\mathrm{S}_I, \gamma_{\mathrm{S},I}\}$ or $\{\gamma_{\overline{\mathrm{S}},I}, \overline{\mathrm{S}}_I\}$ of a sphere with radius r^{probe}. Additionally, it holds that the inside information of the sphere equals the outside information of the molecule, also denoted *molecule outside information*. Thus, an accumulation of probe spheres belongs to Σ.



(a) probe sphere (b) atom sphere

Fig. 6.14: left: the interior (exterior) of a probe sphere is marked by white (dark-gray) points meaning that it bears the outside (inside) information of the molecule; right: an atom sphere as defined in Definition 6.12.

Let the SAS be defined by $\{\Omega^{1:N}_{I,\mathrm{SAS}}, \gamma^{1:N}_{\Omega,I,\mathrm{SAS}}\}$. From these sets we retrieve $\{\Sigma_{I,\mathrm{SAS}}, \Gamma_{I,\mathrm{SAS}}\}$. With the definition of the probe sphere, the grid-based SES can then be calculated by Eq. (6.27).

The Algorithm 6.2 describes the merging of the l^{th} probe sphere with the already merged $(l-1)$

probe spheres, i.e.,

$$\underbrace{\Big(S(\boldsymbol{r}_l, r^{probe})\Big)}_{l^{th}\text{ probe sphere to be merged}} \quad \cup \quad \underbrace{\left(\bigcup_{i=1}^{l-1} S(\boldsymbol{r}_i, r^{probe})\right)}_{(l\text{–}1)\text{ already merged probe spheres}} \quad \cup \quad \Sigma_{I,\mathrm{SAS}}\,.$$

Algorithm 6.2 Merging probe spheres

Require: $\big(\Sigma_{I,\mathrm{SAS}}$ and $\{\mathrm{S}_I^1, \gamma_{\mathrm{S},I}^1\}\big)$ or $\Big(\{\Sigma_I^{(1:l-1)}, \gamma_{\Sigma,I}^{(1:l-1)}\}$ and $\{\mathrm{S}_I^l, \gamma_{\mathrm{S},I}^l\}\Big)$.

Ensure: $\{\Sigma_I^{(1:l)}, \gamma_{\Sigma,I}^{(1:l)}\}$.

if $l = 1$ **then**
  initialize: $\Sigma_I^{(1:l)} = \Sigma_{I,\mathrm{SAS}} \cup \mathrm{S}_I^1$
  $\gamma_{\Sigma,I}^{(1:l)} = \gamma_{\mathrm{S},I}^1$
else
  initialize: $\{\Sigma_I^{(1:l)}, \gamma_{\Sigma,I}^{(1:l)}\} = \{\Sigma_I^{(1:l\text{–}1)}, \gamma_{\Sigma,I}^{(1:l\text{–}1)}\}$
  for all $g_{ijk} \in \mathrm{S}_I^l$ **do**
    insert g_{ijk} in $\Sigma_I^{(1:l)}$
    if $g_{ijk} \in \gamma_{\Sigma,I}^{(1:l)}$ **then**
      remove g_{ijk} from $\gamma_{\Sigma,I}^{(1:l)}$
  for all $g_{ijk} \in \gamma_{\mathrm{S},I}^l$ **do**
    for all $\boldsymbol{e}_d \in \{\pm\boldsymbol{e}_x, \pm\boldsymbol{e}_y, \pm\boldsymbol{e}_z\}$ **do**
      $$\mathrm{e}_{d;ijk}^l \oplus_{\mathrm{P}} \mathrm{e}_{d;ijk}^{(1:l)} = \left\{ \begin{array}{lll} 1 & : \text{insert} & \mathrm{e}_{d;ijk}^l \\ 1^{max} & : \text{update} & \mathrm{e}_{d;ijk}^{(1:l)} \\ 0 & : \text{remove} & \mathrm{e}_{d;ijk}^{(1:l)} \end{array} \right\} \text{ in } \gamma_{\Sigma,I}^{(1:l)}$$
    if $\Big(\Big(N(\mathrm{e}_{ijk}^{\text{non-trivial}}) == 0\Big)$ for $g_{ijk} \in \gamma_{\Sigma,I}^{(1:l)}\Big)$ **then**
      remove g_{ijk} from $\gamma_{\Sigma,I}^{(1:l)}$
return $\{\Sigma_I^{(1:l)}, \gamma_{\Sigma,I}^{(1:l)}\}$

In principle, it is the analogous computation as for the SAS generation (Algorithm 6.1) with the only difference that the probe spheres bear the outside information of the molecule. To generate the SES, we try to maximize Σ of the molecular system starting with the outer domain $\Sigma_{I,\mathrm{SAS}}$. This results in the addition table 6.4 for the merging of probe spheres.

The difference in the action of the $\oplus_A$- and the $\oplus_P$-operator can be seen when comparing Fig. 6.10 and Fig. 6.15: in the former we merged atom spheres to form a molecule whereas in the latter these three spheres play the role of probe spheres. In this case, the sphere P 3 does not have any effect on the configuration as it lies completely outside. In contrast, P 1 and P 2 "dig holes" in the molecule. In this way the outer domain is maximized.

$e^l \oplus_P e^{(1:l)}$		$\overline{0}$	$\overline{1}$	0	1
	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
	$\overline{1}$	$\overline{0}$	$\overline{1}^{max}$	$\overline{1}$	$\overline{0}$

Tab. 6.4: Merging probe spheres with the $\oplus_P$-operator; e^l denotes an edge of the probe sphere (rows) and $e^{(1:l)}$ the corresponding edge of the molecule under construction (column).

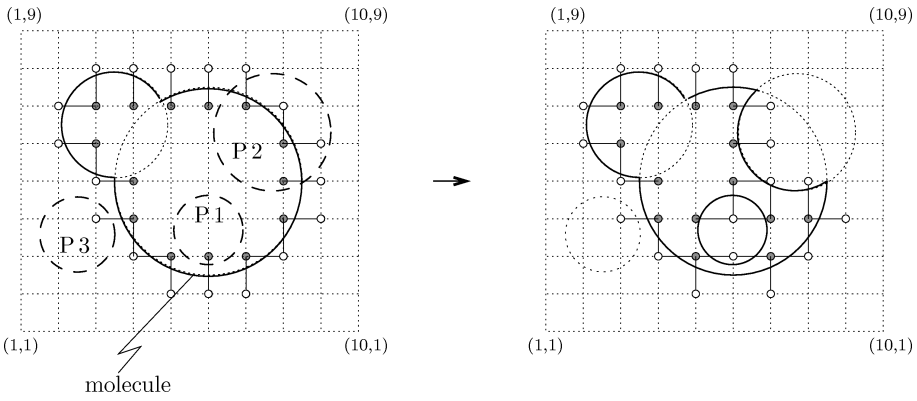


Fig. 6.15: Merging scenarios for the dashed probe spheres (1-3): P 1 lies inside the molecule, P 2 overlaps with the molecular surface and P 3 lies within one grid spacing next to the molecule.

The calculation of the SES of trypsin in a grid with 0.5Å spacing and dimension (64Å×64Å×64Å) takes 193.22 seconds on a machine with four Intel(R) Xeon(R) CPU W3540 @ 2.93GHz processors and 12GB RAM. Fig. 6.16 shows the SES of trypsin generated by this procedure. It is the same cross section of trypsin as the one depicted in Fig. 6.11. The normal vectors drawn as red lines start at the position of SES intersection points. From this, we can see that the surface does not coincide with the VdWS. Further, on the right side of the cross section we see two pockets which are too small for the probe spheres to enter. The SES smoothly closes these pockets.

6.3.4 Final remarks

Comparing the addition tables in Tab. 6.3(b) and Tab. 6.4 it is apparent that under the transformation,

$$(0, 1, \overline{0}, \overline{1}) \rightarrow (\overline{0}, \overline{1}, 0, 1)\,,$$

the $\oplus_P$-operator equals the $\oplus_A$-operator. This corresponds to the fact that the inside information of the spheres, which have to be merged, has always the highest priority regardless of the assigned molecule inside/outside information. This means that the merging procedures for atom and probe spheres are equal, which makes the implementation of the grid-based SAS reusable for the grid-based SES generation.

The complexity of the SES algorithm is mainly determined by the merging procedures of the probe spheres. This is about $\mathcal{O}(M \times n^3)$, where M denotes the number of SAS intersection points

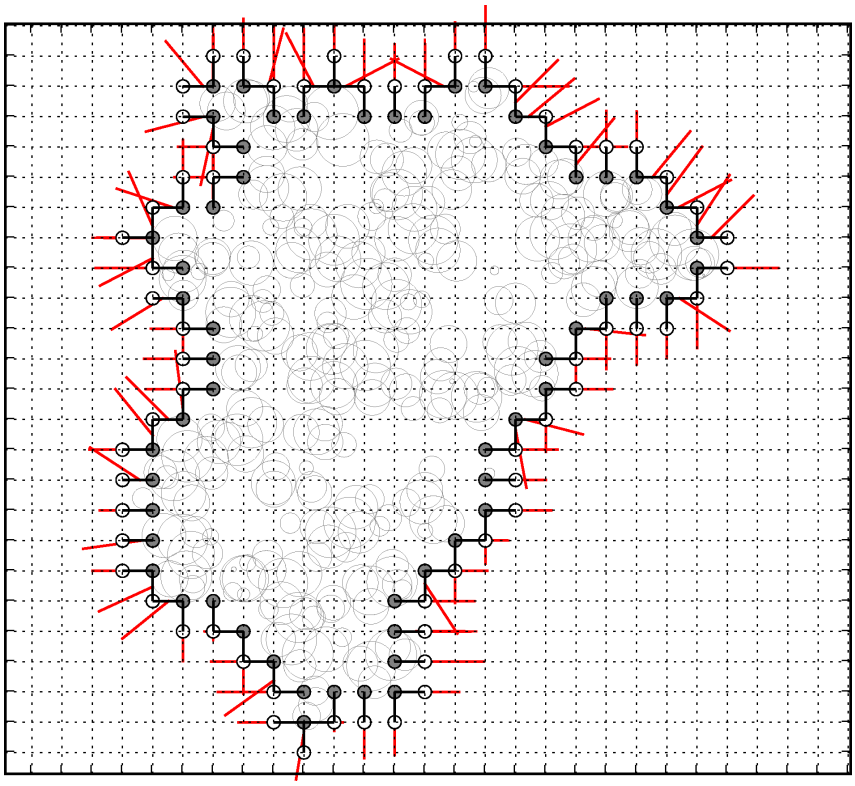


Fig. 6.16: SES constructed by probe spheres hung up at the intersection points of the grid-based SAS, Fig. 6.11(b).

and n^3 the probe grid dimension (rough measure for inside points). Since the number of intersection points M approximately quadruples when the spacing is halved,

$$M_{i+1} \approx 4\, M_i \quad \text{for} \quad h_{i+1} = \frac{h_i}{2}\,, \tag{6.28}$$

the original algorithm becomes very slow for high grid resolutions, $h < 0.125$Å. This *is* however the upper bound of a grid-based resolutions of biomolecules used in finite difference methods.

A first reduction of runtime is possible when taking advantage of the discretization of the potential positions of intersection points: this allows for the precalculation of all possible probe spheres in small reference grids. Further, the SES generation is accelerated by the application of the inner flood-fill algorithm after the complete merging process instead of merging of all inside points of every probe sphere as described in line 2 and lines 6-9 in Algorithm 6.2.

Finally we modified the original algorithm to allow for different resolutions of the SAS and the SES generation. We introduced an user-defined coarse grid resolution h_{SAS} for the SAS and a finer resolution h_{SES} for the reference grids of the probe spheres. The latter is responsible for the SES resolution. If, for instance, h_{SES} equals $\frac{h_{SAS}}{2}$, the number M of SAS intersection points is reduced by approximately a factor 4 (see Eq. (6.28)) and therefore the time required for merging the probe spheres is reduced as well. This procedure lowers the accuracy of the grid-based SES only slightly, because for a grid spacing $h_{SAS} \leq 0.5$ the number of discrete SAS intersection points is still adequate for the merging procedure.

Besides modifications of the generation of the grid-based surface information, which have been discussed before, a great potential to gain efficiency lies in a parallelization of the code. The flood-

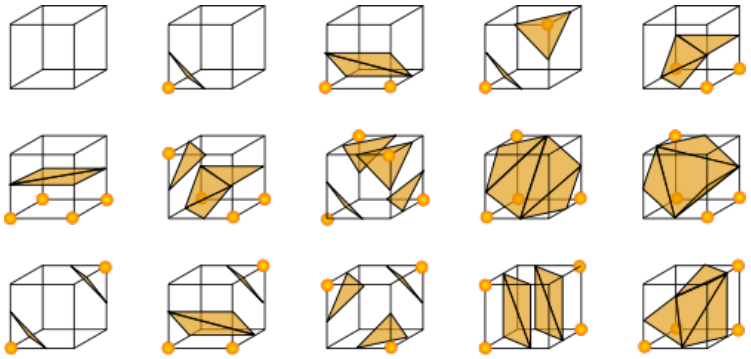


Fig. 6.17: 15 unique cube configurations.

fill algorithms as well as parts of the merging procedures can be parallelized. This will certainly improve the performance.

In summary, the focus of the current implementation was mainly on designing a stable and - most important - accurate generation of the grid-based surface. A detailed study of its possible extensions further improving the efficiency has not been made yet. The extensions which have been proposed and incorporated demonstrate that the implementation can be easily modified to increase its performance.

6.3.5 The marching cubes algorithm

Besides its original purpose to generate the input for the EJIIM solver, we also used the grid-based surface description to calculate a surface triangulation which can be used for the BEM solver after a postprocessing smoothing step. This application relies on the ideas of the marching cubes algorithm.

Marching cubes is an algorithm used in computational geometry for extracting a polygonal mesh of an isosurface from a 3D scalar field f [82]. This means that one searches for the surface where the scalar field takes the iso-value $C \in \mathbb{R}$:

$$(x, y, z) \in \mathbb{R}^3 \quad \text{with} \quad f(x, y, z) = C$$

The marching cubes algorithm proceeds through the scalar field defined in a 3D grid, taking eight neighbor locations at a time and extracting the field values on these eight nodes on this cube. In Fig. 6.17 orange nodes are assigned, for example, a field value $f(x, y, z) < C$. All remaining nodes of the cube have a field value $f(x, y, z) > C$.

In the case of cubes possessing nodes with smaller *and* greater field values, the contour surface has to lie in this cube. Then, we can determine the polygons needed to represent the part of the isosurface that passes through this cube like it is shown in Fig. 6.17 for some possible polygon configurations.

The determination of the correct polygon configuration is done by creating an index to a precalculated array of the 256 possible polygon configurations within the cube, by treating each of the 8 scalar values as a bit in an 8-bit integer. If the value of the scalar field is higher than the iso-value (i.e., it is inside the surface) then the appropriate bit is set to one, while if it is lower (outside), it is set to zero. The resulting value, after all 8 scalars are checked, is the actual index to the polygon configuration entry.

Each vertex of the generated polygons is finally placed on the appropriate position along the

cube's edge by linearly interpolating the two scalar field values that are connected by this edge.

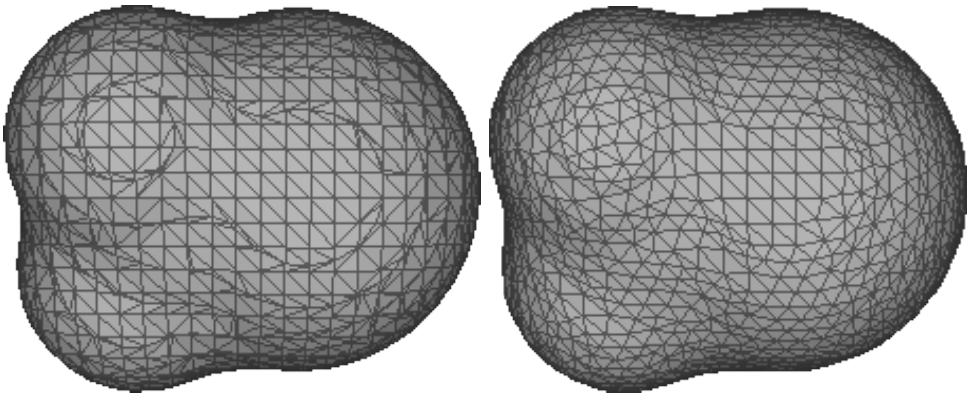


Fig. 6.18: Left: triangulation of a small molecule generated from the EJIIM input data by the marching cubes algorithm; right: Laplace smoothing of the original data.

In order to fit to the original algorithm, we assign a scalar field f composed of zero in $\{\Omega_I, \gamma_{\Omega,I}\}$ and ones in $\{\Sigma_I, \gamma_{\Sigma,I}\}$. If we now search for the contour surface $f(x, y, z) = 0.5$, all the cubes participating in the molecular surface are considered for a non-trivial polygon configuration.

In contrast to the interpolating scheme, which has to be used when determining the contour surface of an arbitrary scalar field, we exactly know the position of the surface on the edges of the grid. Thus, the last step of the original algorithm is modified: we replace the interpolation scheme by directly assigning the coordinates of the intersection point to the position of the polygon.

The left side of Fig. 6.18 illustrates the SES triangulation of a small molecule (5 atoms) generated by the previously described marching cubes algorithm. The intersection points of the molecule with a 3D grid of 0.25 Å spacing are the nodes of the triangulation. As can be seen in Fig. 6.18, the surface triangulation is oriented parallel to the Cartesian grid, for the intersection points lie on

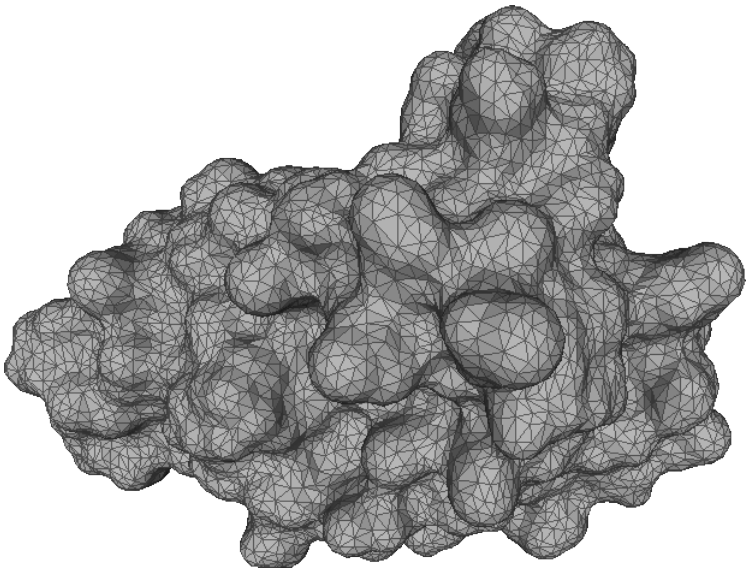


Fig. 6.19: Surface triangulation of a biomolecule generated from the EJIIM input data by the marching cubes algorithm and coarsed by the QECD algorithm.

the Cartesian grid lines. The quality of the triangles is very diverse and the triangle distribution has to be improved before they can serve as an input triangulation for the BEM. The surface visualization program Meshlab [24] offers different tools for remeshing the original triangulation while preserving the surface shape and normal information. The right side of Fig. 6.18 shows the result after applying the Laplacian smoother of Meshlab.

In order to demonstrate that the input generation algorithm handles complex biomolecules as well, Fig. 6.19 illustrates the SES of a CA variant of HIV (GH123) which consists of 2347 atoms. The intersection points have been calculated on a 0.5 Å grid and this results in a grid of roughly 90000 faces and 45000 nodes. Since the number of triangles is too high to serve as input triangulation for the current implementation of the BEM, we coarsened the triangulation with the Quadric Edge Collapse Decimation algorithm (QECD) [44] implemented in Meshlab [24]. The coarsening process removes the "worst" triangles and preserves the surface orientation. With a resulting number of 20000 faces and 10000 nodes this triangulation can be used as input for the BEM.

6.4 Numerical comparison of EJIIM and BEM

Section 6.2 provides the details on the EJIIM implementation of the DSM together with an appropriate boundary value approximation. Further, in Section 6.3 we developed an algorithm to generate accurate surface information for the EJIIM and this completes the requirements to apply the numerical solver.

This section provides the first results of the EJIIM solver for system (6.2) and gives a thorough comparison to electrostatic potentials generated with the existing BEM code. We make several studies for the DSM:

- we check the convergence order of both methods for spherical symmetry (Section 6.4.1)
- we provide a set of small non-trivial molecules (Section 6.4.2). For this test set
  - we analyze the numerical convergence, i.e., how much does the EJIIM solution differ from the BEM (Section 6.4.2.3)
  - we investigate the approximations of the different boundary conditions in the EJIIM (Section 6.4.2.4)
- we compare the runtime of both methods (Section 6.4.2.5)

All calculations presented in this section are based on the standard parameter set ($\varepsilon_\infty = 1.8\varepsilon_0$, $\varepsilon_\Omega = 2\varepsilon_0$, $\varepsilon_\Sigma = 78.5\varepsilon_0$, $\lambda = 3.028$Å)

6.4.1 Convergence towards analytical solution

The convergence studies are done with spherical geometry, as in this case the analytical solution for system (6.2) is available. Additionally, the spherical geometry offers the possibility to generate "exact" input data for both methods.

Ω is a sphere with radius 1Å. We center the sphere at $(0.2, 0.3, 0.01)$ to avoid symmetry effects and locate an unit electron charge $q = e$ at the sphere's center.
For EJIIM, the computational domain is given by $\Sigma \cup \Gamma \cup \Omega = [-5, 5] \times [-5, 5] \times [-5, 5]$ in all cases. The necessary Dirichlet condition along the exterior boundary is given by the analytical solution.

We inspect the convergence with respect to the parameter h, which, in the case of EJIIM, is the usual mesh width. For a triangulation $\mathbb{T}$ with triangles $T \in \mathbb{T}$, h^{bem} is defined as follows (see [121], p. 213):

$$h^{bem} = \max_{T \in \mathbb{T}} h_T \quad , \quad h_T = \sqrt{\text{area}(T)} \tag{6.29}$$

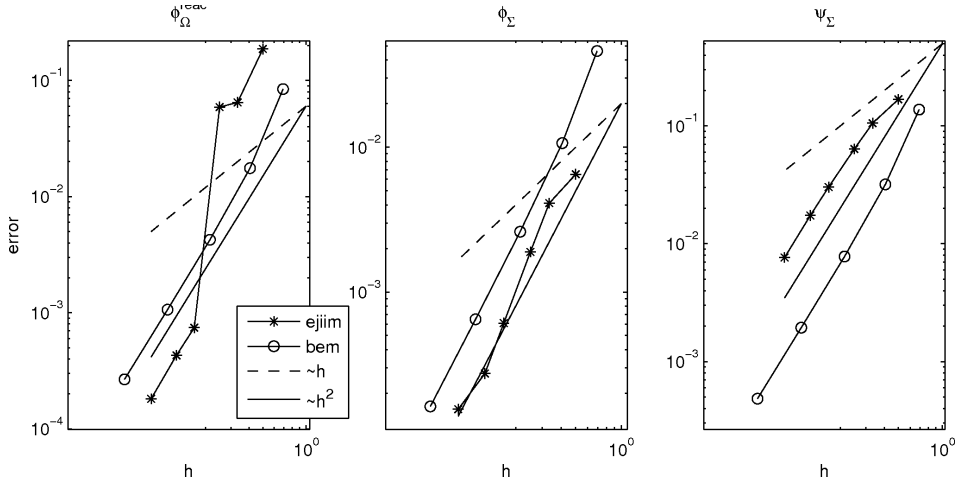


Fig. 6.20: Convergence of the absolute error for the DSM (BEM, EJIIM). Please note remarks in Section 6.4.1.1.

The absolute numerical error has been computed at randomly selected points (50 in Ω, 100 in Σ) for several resolutions (EJIIM) and triangulations (BEM), respectively. Fig. 6.20 confirms that the expected second order convergence for both methods has been achieved.

An important observation can be made from Fig. 6.21, where we plotted the error versus the number of intersection points (EJIIM) and nodes (BEM). For both methods, this is a quantity that is related to the surface resolution and can be directly compared. We see that for the potential ϕ, the EJIIM and the BEM need a similar number of surface nodes to achieve the same quality.

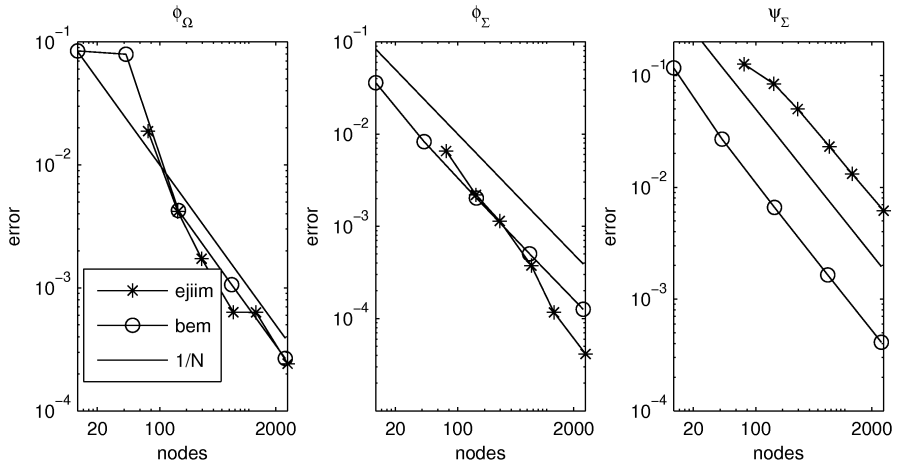


Fig. 6.21: Convergence of EJIIM and BEM with respect to the surface nodes for the DSM (absolute error). Please note remarks in Section 6.4.1.1.

The EJIIM code has also been tested with more general right hand sides and arbitrary given functions in the transmission conditions. In all cases, the second order convergence was confirmed. The BEM implementation of system (6.2) has been tested for a general point charge location inside a Born sphere [52].

6.4.1.1 EJIIM reformulation

Although Figs. 6.20 and 6.21 confirm the theoretically expected second order convergence of the EJIIM and the BEM, the *quantitative* values for the dielectric potential ψ are worse than for the electrostatic potential ϕ when we solve with the EJIIM. The reason lies in the coupling of the fields ϕ_Σ and ψ_Σ in Eq. (6.2e),

$$\left(\Delta - \frac{1}{\lambda^2}\right)\phi_\Sigma + \frac{\kappa^2}{\varepsilon_\infty}\psi_\Sigma = 0 \quad \text{with } \frac{1}{\lambda} = \kappa\sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}\,, \tag{6.30}$$

together with our first choice to handle the set $\{\phi_\Omega^{reac}, \phi_\Sigma\}$ as the transition problem and ψ_Σ as the boundary value problem.

This makes sense from the physical point of view, for we separate the electrostatic from the dielectric potential. However, in practice, this yields the same absolute error e_ϕ for $\{\phi_\Omega^{reac}, \phi_\Sigma\}$ and approximately $e_\psi = \varepsilon_\Sigma\, e_\phi$ for the dielectric field ψ. The reason can be directly seen in Eq. (6.30), which becomes

$$\psi_\Sigma \approx \varepsilon_\Sigma \phi_\Sigma\,,$$

when the differences in the electric potential are small, i.e., $\triangle\phi_\Sigma \approx 0$. This is the case a few Ångstrom away from the biomolecule. Thus, we expect a factor ε_Σ between the error e_ϕ of the electric potential and the error e_ψ of the dielectric potential. Indeed, a closer look at Fig. 6.20 confirms this assumption. In contrast, such a difference in the order of the error cannot be seen for the BEM solution in Fig. 6.20, because every field has its own representation formula *independent* from the other fields.

As we have located the problem in the finite difference formulation, we can try to restate the problem in order to overcome this deficiency. Actually, the following system, which also solves the DSM, can reproduce the outer electric and the dielectric potential, ϕ_Σ and ψ_Σ, in the same order of accuracy:

$$\phi_\Sigma = \frac{1}{\varepsilon_\infty}(\psi_\Sigma - (\varepsilon_\Sigma - \varepsilon_\infty)F_\Sigma) \tag{6.31}$$

$$\begin{aligned}
\triangle\phi_\Omega^{reac} &= 0 && \text{in } \Omega \\
\triangle\psi_\Sigma &= 0 && \text{in } \Sigma \\
\frac{1}{\varepsilon_\infty}\left(\psi_\Sigma - (\varepsilon_\Sigma - \varepsilon_\infty)F_\Sigma\right) - \phi_\Omega &= \phi_{mol} && \text{on } \Gamma \\
\partial_n\psi_\Sigma - \varepsilon_\Omega\partial_n\phi_\Omega &= \varepsilon_\Omega\partial_n\phi_{mol} && \text{on } \Gamma
\end{aligned}$$

$$\begin{aligned}
\left(\triangle - \frac{1}{\lambda^2}\right)F_\Sigma &= -\frac{\kappa^2}{\varepsilon_\infty}\psi_\Sigma && \text{in } \Sigma \\
\partial_n F_\Sigma &= 0 && \text{on } \Gamma
\end{aligned}$$

In fact, we used the possibility given in Eq. (6.31) to switch between the fields ϕ_Σ, ψ_Σ and F_Σ. We solve for the transmission problem in $\{\phi_\Omega^{reac}, \psi_\Sigma\}$ and for the boundary value problem in F_Σ. Assuming an error e for the transmission problem, we directly see that the error in F_Σ scales with $\frac{1}{\varepsilon_\Sigma}$, which finally results in the same accuracy in ϕ_Σ and ψ_Σ. Fig. 6.22 confirms this. The error of the physical relevant potentials, $\{\phi_\Omega^{reac}, \phi_\Sigma\}$ is the same as in the former implementation.

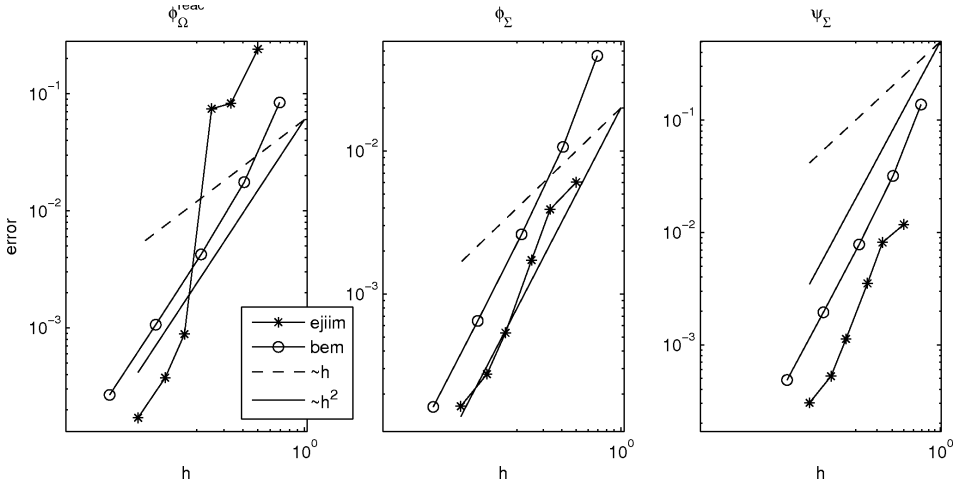


Fig. 6.22: Convergence of the absolute error for the reformulated DSM.

6.4.2 Comparison on molecular geometries

In the following, we analyze the BEM and the EJIIM for non-trivial molecular surfaces. We introduce information about the test set, which is used to study the numerical convergence and the boundary value approximation for the EJIIM. The section closes with a short discussion on the computational cost.

6.4.2.1 Test set

As test set we have selected 6 molecules of the validation suite of MMFF94 [91], three of them with non-vanishing total charge (charged) and three with vanishing total charge (uncharged). Radius and charge assignments have been performed with the MMFF94 implementation in BALL [70]. In Tab. 6.5, we give the numerically relevant parameters for the molecules. The smallest molecule of this set, namely AN05A, is shown in Fig. 6.18 on p. 114.

Tab. 6.5: Overview on the numerical data of the test set, the grid dimension of the EJIIM is the same in all three directions.

Molecule	Grid dimension			Intersections			BEM Elements		
	$h = 0.5$	$h = 0.25$	$h = 0.125$	$h = 0.5$	$h = 0.25$	$h = 0.125$	QECD	Original	BF
charged									
AN05A	31	61	121	482	1904	7556	702	1404	5616
BRMW1	33	65	129	648	2586	10252	642	1286	5144
GUANCH01	39	77	153	1034	4046	16192	1646	3292	13168
uncharged									
FUCMUL	35	69	137	906	3600	14440	1060	2122	8488
FUDPOJ	31	61	121	656	2636	10598	978	1956	7824
ZZZIZA01	39	77	153	1242	4874	19568	1436	2872	11488

All EJIIM computations are performed with three different grids with mesh widths $h \in \{0.5,\ 0.25,\ 0.125\}$. For the BEM, three resolutions are used: the "original" one, generated by [22], roughened by a Quadric Edge Collapse Decimation algorithm (QECD) and refined by a Butterfly Subdivision algorithm (BF) [24, 44].

6.4.2.2 Error measures

Equipped with two different methods to solve the DSM, the question arises how to measure their performance in real world problems, where no analytical solution is available. We decided to stochastically generate point sets, X_Ω and X_Σ, of 100 points inside and outside the molecule, respectively, where both X_Σ and X_Ω lie in the computational box of the EJIIM calculations. The deviation is measured in the following seminorms

$$|u|_\Omega := \max |u(X_\Omega)|\,, \quad |u|_\Sigma := \max |u(X_\Sigma)|\,.$$

Further, for the quantity u we will use the notations u^h: the EJIIM solution with mesh width $h \in \{0.5, 0.25, 0.125\}$, u^{bem}: the BEM solution with the original surface triangulation, u^{qecd}: the BEM solution with QECD-coarsened and u^{bf}: the BEM solution with BF-refined surface.

To compare quantities v and w we will use the error function

$$e(v, w) := \frac{|v - w|_I}{|w|_I}\,, \qquad I \in \{\Omega, \Sigma\}\,. \tag{6.32}$$

Tab. 6.6: Numerical convergence towards the fine grid solution.

	charged			not charged		
Molecule	AN05A	BRMW1	GUANCH01	FUCMUL	FUDPOJ	ZZZIZA01
$e(\phi_\Omega^{0.5}, \phi_\Omega^{0.125})$	2.22e-2	1.58e-2	3.61e-2	5.47e-1	5.66e-1	2.10e-1
$e(\phi_\Omega^{0.25}, \phi_\Omega^{0.125})$	2.65e-3	4.14e-3	5.70e-3	6.18e-2	5.31e-2	5.30e-2
factor	8.39	3.82	6.33	8.87	10.66	3.95
$e(\phi_\Omega^{qecd}, \phi_\Omega^{bf})$	6.28e-3	1.33e-2	9.62e-3	1.94e-1	2.50e-1	1.20e-1
$e(\phi_\Omega^{bem}, \phi_\Omega^{bf})$	4.82e-3	4.97e-3	2.59e-3	6.02e-2	5.82e-2	5.90e-2
factor	1.30	2.67	3.72	3.21	4.29	2.03
$e(\phi_\Sigma^{0.5}, \phi_\Sigma^{0.125})$	9.23e-3	5.43e-3	1.69e-2	2.29e-2	9.43e-2	6.02e-2
$e(\phi_\Sigma^{0.25}, \phi_\Sigma^{0.125})$	1.70e-3	1.10e-3	1.53e-3	1.67e-3	7.67e-3	2.61e-3
factor	5.43	4.92	11.0	13.7	12.3	23.0
$e(\phi_\Sigma^{qecd}, \phi_\Sigma^{bf})$	3.70e-3	1.18e-2	7.72e-3	4.80e-3	1.04e-2	9.58e-3
$e(\phi_\Sigma^{bem}, \phi_\Sigma^{bf})$	2.97e-3	3.38e-3	1.01e-3	1.23e-3	9.75e-3	4.26e-3
factor	1.25	3.50	7.62	3.92	1.06	2.25

6.4.2.3 Discussion of the numerical convergence

Here, we are interested in *numerical convergence* only, thus we use the $(\phi_\Sigma^{bf}, \psi_\Sigma^{bf})$ values as the exterior boundary condition for the EJIIM computation. Results generated with this boundary condition can be found in Tab. 6.6 and Tab. 6.7 for the physically relevant electric potential ϕ.

From the EJIIM as well as from the BEM results listed in Tab. 6.6, we cannot directly deduce a specific convergence rate. This is due to the non-trivial input data. Additionally, the randomly chosen point sets, on which the error function is evaluated, only allow a rough interpretation: the outer electric potential of uncharged molecules, ϕ_Σ, has higher absolute errors and faster EJIIM convergence rates than those of charged molecules on average. The reason lies in the fast decrease of the potential outside, which is better mapped for higher grid resolutions.

With Tab. 6.6 we demonstrate that both methods really converge to the fine grid solutions, as the factor is always greater than 1.

In Tab. 6.7, we study the EJIIM solution versus the BEM solution. The difference between the finest grid solutions, $\phi_{\Omega,\Sigma}^{0.125}$ for EJIIM and $\phi_{\Omega,\Sigma}^{bf}$ for BEM, rounded upwards to one significant digit, is used as a rough estimate for the numerical error later when approximating the boundary condition. It is important to mention that in Tab. 6.7 we do not consider the numerical error, that is, the difference between the exact and numerical solutions. Instead, we are looking for the difference between two solutions, where $\phi_{\Omega,\Sigma}^{bf}$ has been used as the reference. Thus, stagnating difference, like for BRMW1, is legitimate.

Tab. 6.7: Convergence towards BEM results.

	charged			not charged		
Molecule	AN05A	BRMW1	GUANCH01	FUCMUL	FUDPOJ	ZZZIZA01
$e(\phi_\Omega^{0.5}, \phi_\Omega^{bf})$	2.05e-2	1.35e-2	3.56e-2	5.75e-1	5.24e-1	2.05e-1
$e(\phi_\Omega^{0.25}, \phi_\Omega^{bf})$	2.70e-3	3.75e-3	5.38e-3	8.71e-2	6.79e-2	4.77e-2
$e(\phi_\Omega^{0.125}, \phi_\Omega^{bf})$	2.36e-3	4.26e-3	1.54e-3	3.91e-2	3.08e-2	1.87e-2
error estimate	3e-3	5e-3	2e-3	4e-2	4e-2	2e-2
$e(\phi_\Sigma^{0.5}, \phi_\Sigma^{bf})$	8.69e-3	7.68e-3	1.65e-2	2.29e-2	9.39e-2	6.02e-2
$e(\phi_\Sigma^{0.25}, \phi_\Sigma^{bf})$	1.62e-3	3.63e-3	9.98e-4	1.67e-3	7.15e-3	2.64e-3
$e(\phi_\Sigma^{0.125}, \phi_\Sigma^{bf})$	1.27e-3	2.67e-3	1.04e-3	1.30e-3	2.72e-3	1.06e-3
error estimate	2e-3	3e-3	2e-3	2e-3	3e-3	2e-3

6.4.2.4 Discussion of the boundary approximations

In Section 6.2.4.1 on p. 94 ff. we proposed boundary conditions. To check the quality of these boundary conditions, we use again our test set. All calculations in this section are done on the finest mesh with $h = 0.125$. In addition, all calculations have been carried out on two different box dimensions with equal resolution in order to investigate the influence of the approximations: 4[Å] and 6[Å] away from the molecular surface, respectively, see Fig. 6.23 for an illustration.

For the *reference solution* u^{ref}, where $u \in \{\phi_\Omega, \phi_\Sigma, \psi_\Sigma\}$, the exterior boundary condition is provided by the BEM calculation on the refined mesh. In Tab. 6.8 the values are computed as $\tilde{e}(u) := \|u - u^{ref}\|_\infty / \|u^{ref}\|_\infty$, $u \in \{\phi_\Omega, \phi_\Sigma\}$ with different boundary conditions for $(\phi_\Sigma, \psi_\Sigma)$.

Tab. 6.8: Approximation of the exterior boundary condition; hom: homogeneous, edp: effective dipole, approx: approximated, e: estimate of the numerical error.

$\tilde{e}(\phi_\Omega)$							
	4[A] Box			6[A] Box			
Molecule	hom	edp	approx	hom	edp	approx	e
AN05A	1.69e-1	2.59e-2	2.51e-2	4.95e-2	5.54e-3	5.43e-3	3e-3
BRMW1	2.25e-1	6.73e-2	4.04e-2	5.76e-2	1.19e-2	7.92e-3	5e-3
GUANCH01	1.01e-1	5.86e-2	1.29e-2	3.19e-2	9.79e-3	2.63e-3	2e-3
FUCMUL	7.74e-1	7.74e-1	6.38e-2	1.10e-1	1.10e-1	1.09e-2	4e-2
FUDPOJ	3.98e-1	2.97e-1	4.36e-2	6.16e-2	3.63e-2	8.35e-3	4e-2
ZZZIZA01	9.27e-1	9.27e-1	8.28e-2	1.35e-1	1.35e-1	1.15e-2	2e-2

$\tilde{e}(\phi_\Sigma)$							
	4[A] Box			6[A] Box			
Molecule	hom	edp	approx	hom	edp	approx	e
AN05A	1.90e-1	3.17e-2	2.95e-2	7.54e-2	1.16e-2	1.09e-2	2e-3
BRMW1	3.04e-1	1.13e-1	5.74e-2	1.16e-1	3.63e-2	2.07e-2	3e-3
GUANCH01	1.38e-1	1.06e-1	1.96e-2	5.77e-2	3.56e-2	7.51e-3	2e-3
FUCMUL	9.50e-2	9.50e-2	6.35e-3	2.86e-2	2.86e-2	2.06e-3	2e-3
FUDPOJ	7.84e-2	8.07e-2	5.91e-3	2.46e-2	2.16e-2	2.02e-3	3e-3
ZZZIZA01	2.25e-1	2.25e-1	2.10e-2	7.05e-2	7.05e-2	6.41e-3	2e-3

Homogeneous, approximated, and effective dipole conditions A closer look at Tab. 6.8 suggests that the approximated boundary condition provides the best approximation: here, $\tilde{e}$, which accounts for the error made by the boundary value approximation, is always of the same order as the numerical error, which was extracted from Tab. 6.7.

The effective dipole and even the homogeneous ansatz can reasonably reproduce the BEM results. Note that for all molecules of our test set with vanishing total charge, the effective dipole approximation does not lead to better results than the homogeneous boundary condition. This originates from the spatial charge distribution: for vanishing effective monopole and dipole (mirror symmetry of the charge distribution for example) the effective dipole approximation is in fact a homogeneous boundary condition and therefore leads to the same accuracy. In contrast, for the charged molecules of our test set, the effective approximation is comparable with the approximated boundary condition when the 6[Å] computational box is taken.

In Fig. 6.23, we illustrate the behavior of the approximated boundary conditions for an (un)charged molecule, (FUDPOJ) GUANCH01. The difference between the numerical solution of system (6.2) and Eq. (6.19), evaluated on Σ, is additionally plotted in Fig. 6.23 and supports our assumption that the influence of the dielectric boundary is negligible for the outer field solutions a few Ångstrom from the molecule when $\frac{\varepsilon_\infty}{\varepsilon_\Omega} \approx 1$.

Focusing condition The focusing boundary condition is often used for high-quality calculations, where the boundary values of the coarse grid calculations are usually fast generated ap-

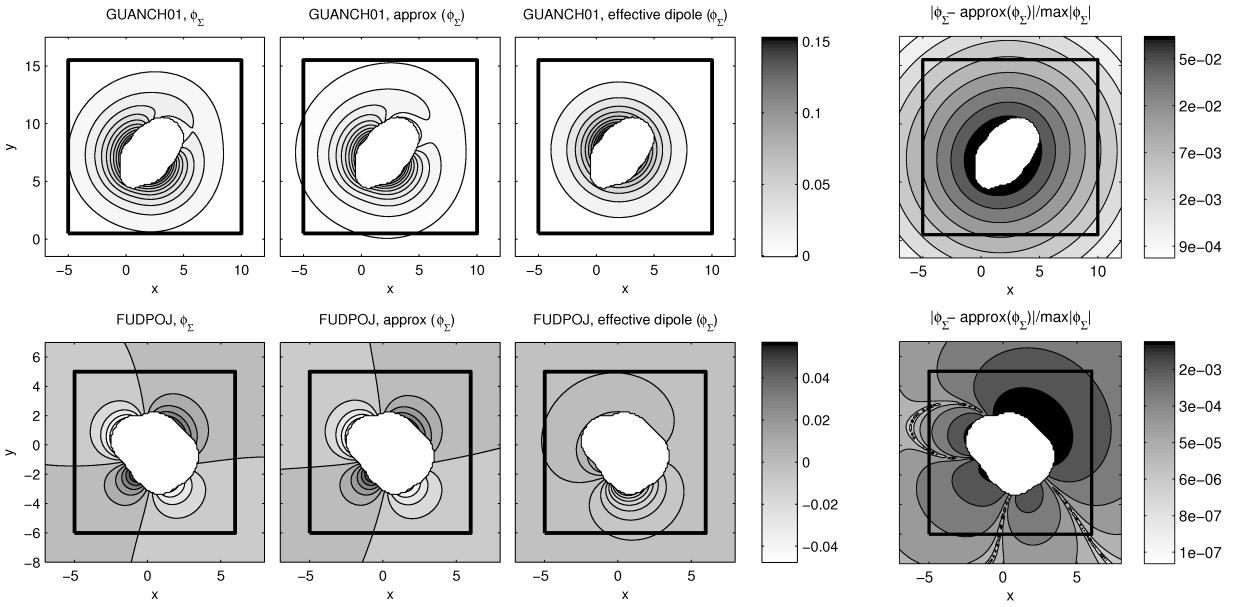


Fig. 6.23: Reference solution ϕ_Σ, approximated and effective dipole potentials. Cuts along the plane $z = z^*$, where (x^*, y^*, z^*) is the center of the computational box. The black border shows the 4[Å] box. Contour lines of the difference are selected logarithmically to visualize also the small values.

proximations. In our calculations, the effective boundary condition turns out to be a good choice. Overall, the focusing method has roughly the same accuracy as the dipole boundary condition when the 4[Å] computational box is taken.

We conclude that the derived boundary value approximations do not reduce the quality of the solution strongly. The special choice of the boundary condition depends on the demand on accuracy and efficiency.

6.4.2.5 Runtime analysis

The EJIIM approach was implemented using Matlab version R2007b, the BEM was implemented in `C` and uses the ATLAS library [139].

To give a rough impression on the computational costs, we compared both methods on a machine with eight Intel(R) Xeon(R) CPU X5355 @ 2.66GHz processors and 8GB RAM. In Tab. 6.9, we list the runtime necessary for the EJIIM calculation of AN05A (small box) with homogeneous boundary condition and those of the BEM calculation to generate the Cauchy data.

The runtime and iteration counts for other molecules are similar. In order to relate the results to the numerical quality and to the surface information, we refer to Tab. 6.8 and Tab. 6.5, respectively.

Of course, the values listed in Tab. 6.9 does not allow us to answer the question if one or another method is faster or requires less memory. These factors highly depend on the programming language, particular implementation, and the computational system. In addition, the output of EJIIM (full grid solution) differs from the BEM result (certain data along the surface). However, we see that the runtime of both methods are of approximately the same order.

Tab. 6.9: Computational costs for AN05A.

EJIIM					BEM	
type	runtime (sec) matrix setup	runtime (sec) solving	outer BiCGSTAB iterations (10^{-8} stopping tolerance)	inner BiCGSTAB iterations (10^{-10} stopping tolerance)	type	runtime (sec) solving Cauchy Data
0.5	5	7	6.5	4	*qecd*	3
0.25	41	78	7.5	4	*bem*	20
0.125	1023	1009	8.5	4	*bf*	821

6.5 Conclusion

In this chapter we developed new numerical tools to investigate nonlocal electrostatic models. We found that the EJIIM is a competitive alternative to the BEM to solve the elliptic differential equations [138]. It is comparable with the BEM in terms of accuracy and runtime. What sets EJIIM apart and what finally makes it the preferable numerical method are the following aspects:

- A partial differential equation system (PDES) can be fast and efficiently implemented using the EJIIM. Further, the method is fully automatable. This makes it applicable in biomolecular software and opens the research on nonlocal electrostatics to a broad community.
- An extension to nonlinear PDES is possible in the EJIIM. This is necessary to incorporate the nonlinear Poisson-Boltzmann equation into the PDES of nonlocal electrostatics. In the BEM, solving a nonlinear equation is much harder to realize.
- The required amount of memory allocated when solving the PDES depends on the size of the biomolecule in consideration. In contrast to the BEM, the system matrices in the EJIIM are sparse, so that the memory issue does not pose a serious problem.
- The algebraic equations can be solved by multigrid methods which are among the fastest solution techniques known today.

In the following two chapters we focus on an investigation of various nonlocal models and its application to biomolecules. Within these chapters, we successfully demonstrate that the EJIIM solver satisfies the listed features.

Chapter 7

Model comparison and applications

In Chapter 4 we have derived two vectorial models to describe the nonlocal effects of the water network, the Newton Vector Model (NVM) and the Dirichlet Vector Model (DVM), where for the NVM a BEM implementation already exists. Based on these vectorial models, we derived two scalar models in Chapter 5, the Dirichlet Scalar Model (DSM) and the Newton Scalar Model (NSM).

In the previous chapter, we discussed in detail two numerical methods, BEM and EJIIM, for solving (linear) differential equation systems. On the basis of the DSM, we gave evidence that these methods are appropriate to solve (non)local electrostatic problems. This statement holds not only for the DSM, but also for the other nonlocal models, NVM and NSM, as well as for the local model (LM) as Fig. 7.1 clearly demonstrates by means of the analytically solvable Born model. The NVM solver, which is used for all the following studies, has been implemented and kindly provided by C. Fasel [39].

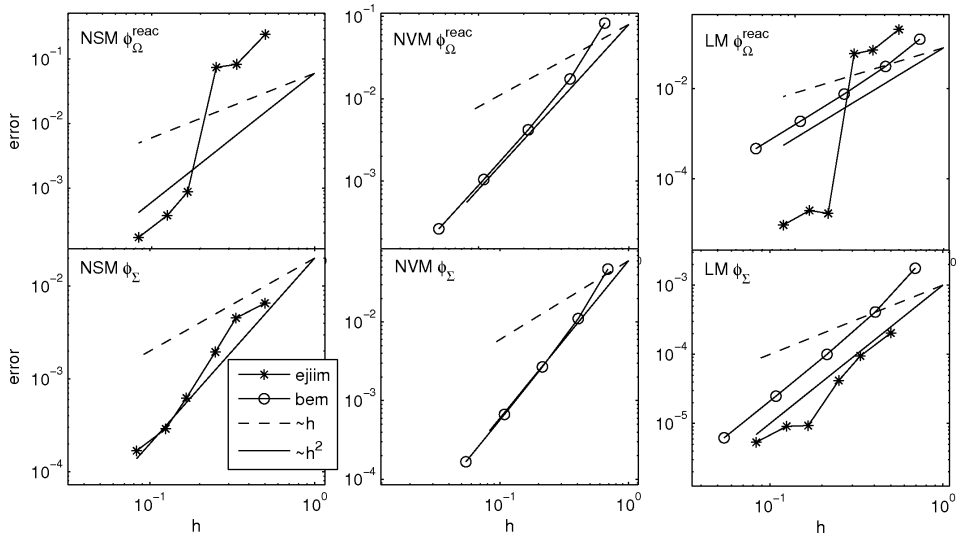


Fig. 7.1: Second order convergence of the NSM, the NVM and the LM in the new EJIIM and in the available BEM implementations.

The second order convergence can be seen for both numerical methods in all models. The unexpected "jump" of the error $e(\phi_\Omega^{reac})$ in ϕ_Ω^{reac} that appears in the EJIIM calculations in Fig. 7.1 originates from a "bad" choice of sample points, meaning that a large number of the randomly chosen points lie near Γ, which becomes better resolved when the resolution gets finer. This is in line with the behavior of $e(\phi_\Omega^{reac})$ that turns into the second order dependence for $h \leq 0.25$.

Having provided different nonlocal models, we are now interested in analyzing the differences of the electrostatic potential ϕ of various molecules. First, this study aims at clarifying whether the approximations assumed in the scalar models preserve the nonlocal features, which we found

for the analytically solvable systems, in non-trivial applications, as well. Second, for an advanced model development and modification, a comparison of the various nonlocal models is necessary in order to estimate the effects of predefined boundary conditions of the correlation field $\boldsymbol{F}$. Thus, the study of the derived models is a good starting point, as all of them force a different behavior of $\boldsymbol{F}$ on the molecular surface. Third, another focus of our research is to find an appropriate formulation of the nonlocal dielectric response of water, which can be solved efficiently and fast. This is the basis to successfully incorporate a solver for nonlocal electrostatics into a biomolecular software package such as BALL and in this way to offer a broad research community the possibility to apply nonlocal as well as local electrostatic models. A comparison of the various models is required to find an answer to this question.

In Section 7.1, a comparison of the scalar models and the NVM on the test set clarifies the effects of the approximations and assumptions we made for the scalar formulations. Further, a comparison of the local and the nonlocal electrostatic potential on realistic biomolecules is given in Section 7.2. On the basis of these investigations, in Section 7.3, we conclude that all the nonlocal models, which we carefully derived in the last sections, are adequate for modeling nonlocal electrostatics.

7.1 Comparison on the test set

For the first study of the nonlocal models we use the test set which we introduced in Section 6.4.2. Since the molecules are small and carry only a few partial charges that sum up to $0e$, $1e$ or $-1e$, we expect fundamental differences between the models to be directly obvious.

The analysis is done using the standard parameter set ($\varepsilon_\Omega = 2\varepsilon_0$, $\varepsilon_\Omega = 78.5\varepsilon_0$, $\varepsilon_\infty = 1.8\varepsilon_0$, $\lambda = 3.028$Å).

7.1.1 Solvation energy

The solvation energy is an important quantity in biomolecular studies, because it is experimentally accessible. For the spherically symmetric case, we have seen in Section 3.3.3.4 on p. 48 that the nonlocal models predict very similar solvation energies and that the nonlocal estimation is up to an order of magnitude lower than the energy approximation of the local model.

There are various aspects which complicate a straight forward interpretation of the electrostatic contribution to the solvation energy of realistic molecules: we already mentioned the use of the atomic radii as fit parameters for solvation models in biomolecular research. An independent estimation of the radii as it is possible for monoatomic ions is difficult already for molecules consisting of a few atoms. Parameter sets including radius information, which have been optimized for a local macroscopic dielectric response, of course bias the solvation energy estimation in the nonlocal framework.

For all models, the atomic radii and charge information are taken from the MMFF94 parameter set. As these radii are optimized for local electrostatics, an interpretation of the electrostatic contribution to the solvation energies predicted in the various electrostatic models is possible only in a qualitative way. The electrostatic solvation energies given in Tab. 7.1 describe the difference in the electrostatic field energy when moving the molecule from vacuum to water. The energies are calculated with the formula (3.2) given in Section 3.1.4

For the LM and the DSM, a BEM and an EJIIM implementation are available. These two different numerical solvers reveal electrostatic solvation energies in excellent agreement as it is shown in Tab. 7.1.

The local model predicts higher solvation energies for both, the charged (AN05A, BRMW1, GUANCH01) and the uncharged molecules (FUCMUL, FUDPOJ, ZZZIZA01). This is due to the great difference between the dielectric response in vacuum and in water assumed in the local model.

Tab. 7.1: Solvation energy in units of $[KJ/mol]$ for the molecules in the test set (EJIIM: 0.25 spacing, BEM: original triangulation.

	LM (BEM)	LM (EJIIM)	DSM (BEM)	DSM (EJIIM)	NSM (EJIIM)	NVM (BEM)
AN05A	-292.94	-292.26	-199.08	-199.62	-199.09	-203.58
BRMW1	-250.08	-249.28	-175.19	-175.81	-175.57	-183.03
GUANCH01	-253.95	-252.98	-166.59	-166.82	-165.84	-157.42
FUCMUL	-20.27	-20.11	-6.58	-6.59	-6.32	-6.93
FUDPOJ	-30.44	-30.00	-9.38	-9.41	-9.07	-9.72
ZZZIZA01	-16.96	-16.86	-6.93	-6.95	-6.72	-7.39

The strong macroscopic polarization of $78.5\varepsilon_0$ in water lowers the field energy, because the dipolar medium effectively screens the electric field of the fixed charges of the biomolecule.

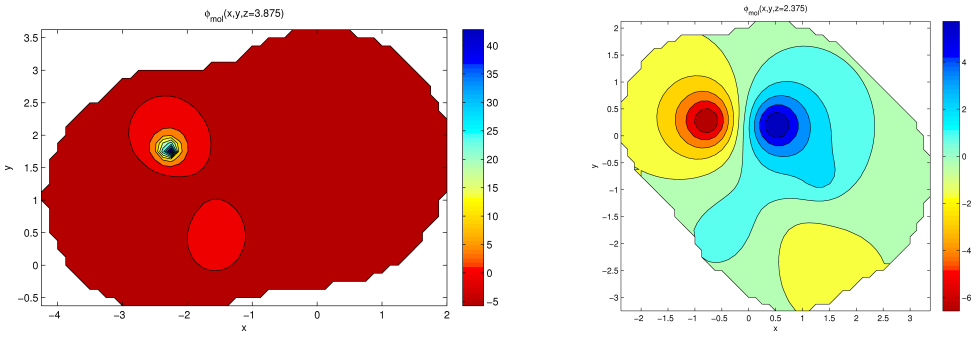
In contrast, the nonlocal models capture the change in the dielectric response due to the water network. The water molecules with their dipolar character do not react individually to the external field, but try to keep their hydrogen bonds by reorienting the molecules in a favorable position. In summary, this causes a decrease in the macroscopic response. Thus, the change in field energy, when moving the molecule from vacuum to the water network, is smaller compared to the local prediction.

The differences in the solvation energy between the scalar nonlocal models and the vector model are very small. This is, at first glance, surprising because the DSM specifies a different behavior on the molecular surface than the NVM or the NSM. However, we learned in Remark 4.4 on p. 63 that the field energy variation generally consists of boundary and volume integrals. In order to obtain a closed formula for the field energy, we approximated the latter by the volume integral. The additional contributions, i.e., those of the boundary integrals, have been interpreted as first shell effects such as, for instance, the reformation of the hydrogen bond network when building the cavity of the solute. The contribution which we take into account at the moment is due to electrostatic effects, only. The very good agreement between the energy predictions therefore demonstrates that the *electrostatic* contribution of the different nonlocal models is similar. With respect to the scalar models, this also means that the approximations which we accepted in Chapter 5 are adequate to find physically reasonable electrostatic energy terms.

7.1.2 Electrostatic potential

In this section, we study the (non)local models by an investigation of the electrostatic potential ϕ. For the discussion, we take two representatives out of the test set: the charged and the uncharged molecule, BRMW1 and FUDPOJ, respectively. The results for the remaining (un)charged molecules have been analyzed in the same way and reveal a similar spatial behavior.

We study the differences of the various models on the basis of contour plots, where a cross section through the molecule is considered. Of course, the results, which can be seen for the chosen cross section, highly depend on the charge pattern next to this plane. This charge pattern can simply be visualized by the molecular field ϕ_{mol} defined in Section 3.1.4, because it accounts for the point charges by their corresponding $\frac{1}{r}$ potential in space. In Fig. 7.2, the molecular field ϕ_{mol} is plotted for the cross sections that we will focus on in detail in Sections 7.1.2.1 and 7.1.2.2.



(a) BRMW1 (b) FUDPOJ

Fig. 7.2: The molecular potentials of the molecules under study through a z-cross section.

7.1.2.1 BRMW1

As we can see from the almost homogeneous red colored region in Fig. 7.2(a), the charge pattern of BRMW1 is globally dominated by the total charge $Q=-e$ of the molecule. However, the molecule possesses a positive partial charge at position $(-2.29, 1.78, 3.86)$, which, in addition, causes a high positive peak of the molecular potential.

Fig. 7.3 shows the electrostatic potential inside the molecule for all analyzed models. Here, the first row depicts the ϕ, it is clearly dominated by the molecular potential given in Fig. 7.3. The second row shows the reaction field inside the molecule. In the nonlocal models, the reaction field is approximately $1[V]$ almost everywhere with slight variations on the longitudinal axis. This reminds of the constant reaction field in the case of the Born model (Section 3.3.3) and indeed, if we assume the molecule to have roughly a radius $r = 3$Å and a total charge $Q=-e$, we find

$$\phi_\Omega^{reac} \approx \frac{c}{r}\left(\frac{1}{\varepsilon_\Sigma} - \frac{1}{\varepsilon_\Omega} + \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma}\right)\frac{1}{1+r/\lambda}\right) \approx 1[V]\,, \tag{7.1}$$

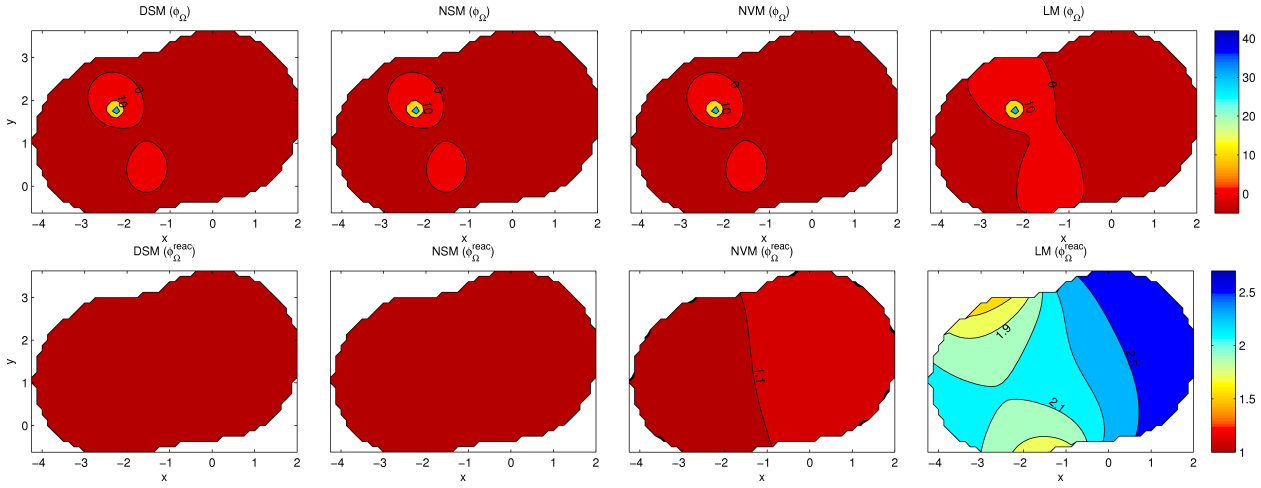


Fig. 7.3: Contour lines of the inner electrostatic potential in [V] of BRMW1. The first row and the second row shows the complete inner potential ϕ_Ω and the reaction field potential ϕ_Ω^{reac}, respectively.

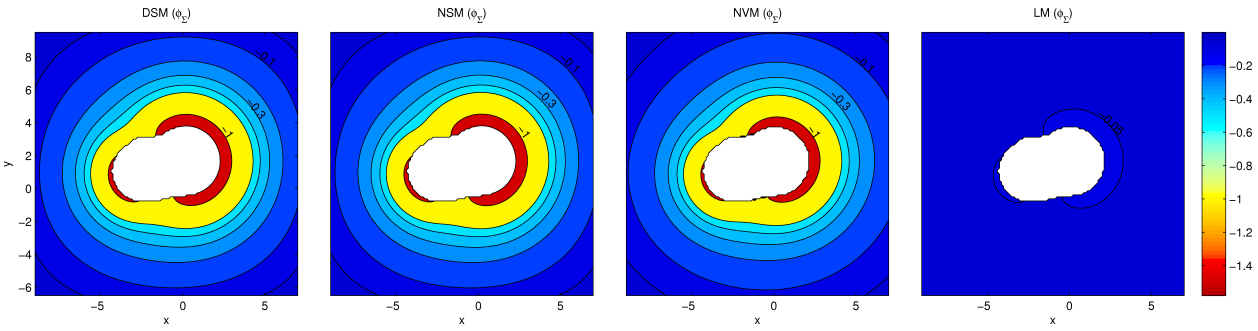


Fig. 7.4: Contour lines of ϕ_Σ in [V] of BRMW1.

where $c = \frac{-e}{4\pi\, 10^{-10}}\,[C]$. This means that the reaction field can be characterized by an effective Born sphere.

In contrast, the reaction field of the local model shown in the second row on the right of Fig. 7.3 has visible variations on the longitudinal axis meaning that the partial charges locally have an impact on the overall behavior. On average, the Born approximation analogous to Eq. (7.1) for the local reaction field reveals the correct order of magnitude, as well:

$$\phi_\Omega^{reac} \approx \frac{c}{r}\left(\frac{1}{\varepsilon_\Sigma} - \frac{1}{\varepsilon_\Omega}\right) \approx 2.27[V]$$

In Fig. 7.4 we plotted the outer electric potential. All models develop the same qualitative behavior, which can be seen on the contour lines. The contour lines become more and more spherical, which is in accordance with the increasing influence of the total charge immersed in an almost spherical molecule.

The potentials in all the nonlocal models exhibit almost the same behavior. Comparing the extent of the red region which starts at the dielectric boundary Γ, we recognize that this region is smaller in the NVM, meaning that the electrostatic potential decreases slightly faster compared to the nonlocal scalar models. From Section 4.3.1, we know that the NVM treats the water network as unperturbed by the solute. Thus, the water network communicates near the molecular surface in the same way as in regions farther away. This results in a higher shielding as, for instance, in the case of the DSM, where the boundary condition on the surface forces the correlation field to be zero.

In contrast to the nonlocal models, which predict the electric potential to reach a few Ångstrom away from the molecule, the local potential decays very fast so that it seems to be already zero on the chosen color scale.

The shape of the outer electrostatic potential and the inner reaction field have to be interpreted together: let us first assume the charges of the molecules to be directly immersed in the solvent, i.e., without any variations in the dielectric response. Then, the system is characterized by the fundamental solutions given in Theorem 3.3 for the nonlocal case and in Theorem 2.1.2 for the local case. Further, this implies that the reaction field is zero. “Inserting” the dielectric boundary corresponds to a non-vanishing reaction field. The magnitude and the variations of the reaction field directly depend on the magnitude of the change in the dielectric response: the stronger the change of the latter, the stronger are the variations of the reaction field. In the local setting, the dielectric response on Γ changes from $\varepsilon_\Omega = 2\varepsilon_0$ to $\varepsilon_\Sigma = 78.5\varepsilon_0$, resulting in noticeable variations of the reaction field and a highly screened outer electrostatic field. In contrast, in the nonlocal models, such a strong difference is not realized on Γ, because the hydrogen bond network hinders the orientational polarization. The dielectric transition is almost smooth. Thus, the reaction field

has small variations, whereas the outer electrostatic potential resembles a smooth extension of the inner electrostatic potential, implying that on Γ it is very similar to the fundamental solution of the local model with dielectric response ε_Ω.

7.1.2.2 FUDPOJ

FUDPOJ is an overall uncharged molecule. Its charge distribution can be imagined as an electrostatic quadrupole consisting of two positive and two negative charges, arranged on the corners of a square. In Fig. 7.2 on p. 128 we can identify this charge pattern, where one of the dipole pairs lies closer to the considered xy-plane.

In Fig. 7.5(b), we plotted the reaction field inside and the electrostatic potential outside of the molecule for all (non)local models. As already discussed in the last paragraph, we clearly see the tendency of the reaction field to be almost constant in the nonlocal model. Because of the vanishing total charge of FUDPOJ, the reaction field therefore varies slightly around zero. In contrast, the local reaction field is higher in value and mimics the charge distribution.

On the lower side of the molecule, the NVM exhibits contrary variations compared to the LM, i.e., the reaction field has a slight positive value in the region of positive charge and a negative

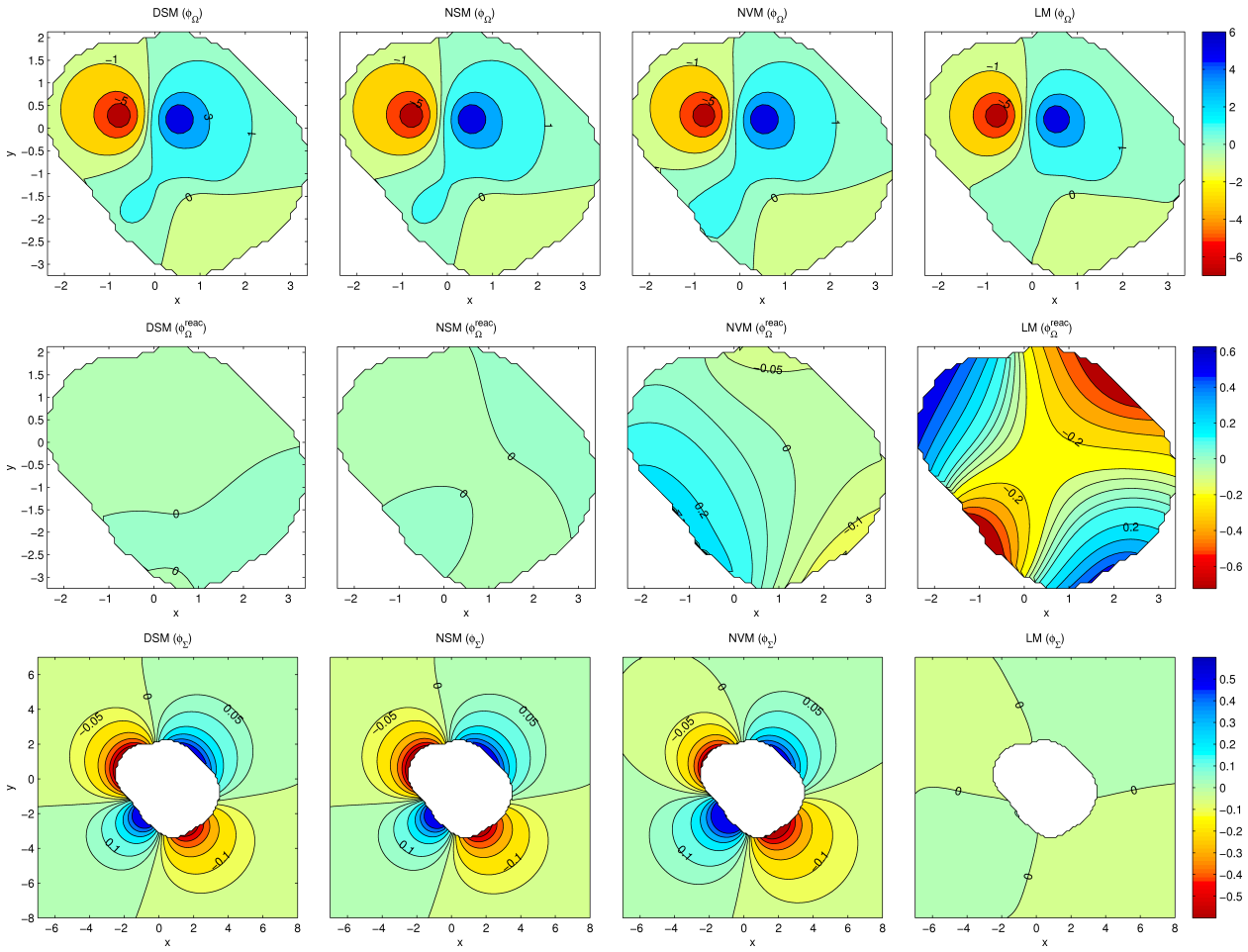


Fig. 7.5: Contour lines of the electrostatic potential (in [V]) of FUDPOJ. The first row and the second row shows the complete inner potential and the reaction field potential, respectively. The third row shows the outer potential.

value where the negative charge is located. This can be explained by noting that the charges which are responsible for the electrostatics within this region lie next to the dielectric boundary, Γ as can be seen in Fig. 7.2(b) by the broadening of the blue (positive charge) and yellow (negative charge) regions towards Γ. Recalling the reaction field for a charge q_{loc} located in the center of the sphere, we see that for $a<\lambda$, ϕ_Ω^{reac} can change the sign

$$\phi_\Omega^{reac} \approx \lim_{a\ll\lambda} \left\{ \frac{c}{a} \left(\underbrace{\frac{1}{\varepsilon_\Sigma} - \frac{1}{\varepsilon_\Omega}}_{<0} + \underbrace{\left[\left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma} \right) \frac{1}{1+a/\lambda} \right]}_{\text{for } \lim_{a\ll\lambda} \Rightarrow (\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma})} \right) \right\} = \frac{c}{a} \left(\underbrace{\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Omega}}_{>0} \right), \tag{7.2}$$

with $c = \frac{q_{\text{loc}}}{4\pi\, 10^{-10}}\,[C]$. We can understand this change in the reaction field when remembering its physical meaning: in Section 2.2.2 as well as in Remark 3.1 on p. 28, we stated that the reaction field can be interpreted as a field of induced surfaces charges created by the dielectric boundary, i.e., induced by the change in the polarization $\boldsymbol{P}$.

$$\sigma_{ind} \overset{\text{on}\,\Gamma}{=} -\boldsymbol{n} \cdot (\boldsymbol{P}_\Sigma - \boldsymbol{P}_\Omega)$$

Comparing the LM and the NVM, we therefore interpret the change in the sign of ϕ_Ω^{reac} to originate from a change in the induced surface charges: in the local model, the solvent has a predefined, stronger dielectric response than the biomolecule, always resulting in induced surface charges opposite in sign compared to the neighboring fixed partial charges of the molecule. In contrast, in the NVM, the solvent varies in its response. In particular, as discussed before, it has a smaller response compared to ε_Ω and therefore induced surface charges of the same sign as the molecular charges when the latter are very close to the dielectric boundary. For the two scalar models, this is in principle also possible, as Eq. (7.2) holds for all nonlocal models. However, the DSM and the NSM seem to be dominated by the vanishing total charge inside Ω.

The star-like shape of the outer electrostatic potential can directly be understood in terms of the charge distribution described above. The nonlocal potential once again reaches far into space in contrast to the local prediction. Comparing the contour lines in the nonlocal models, we see that they differ slightly in their spatial extension. The NVM predicts a higher potential on the lower side, which is a consequence of the near-charge-effect discussed before. The differences in the two nonlocal scalar models, the DSM and the NSM, are marginal.

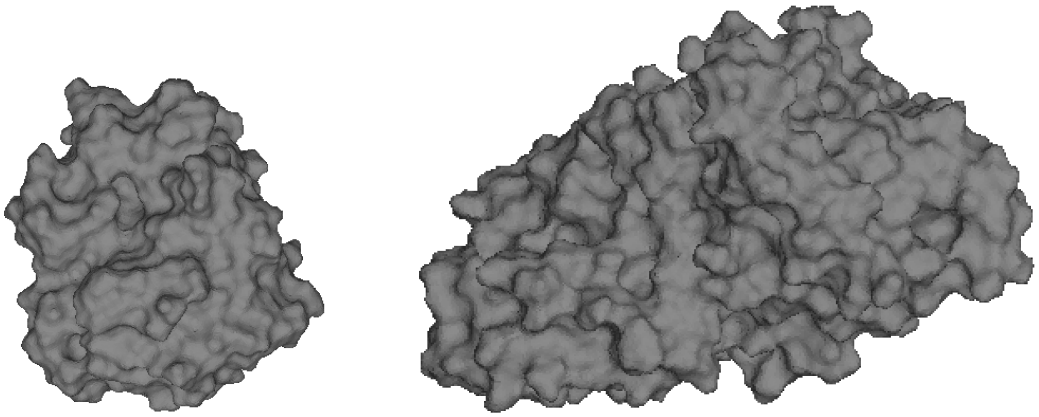
With the comparison on the small molecules, we could demonstrate that all the proposed and implemented nonlocal models exhibit qualitatively the same behavior in space. This behavior is in agreement with what has been expected: in contrast to the local macroscopic model, the nonlocal models predict a potential, which is higher in magnitude outside, and, in return, smaller in magnitude inside the molecule.

Since both scalar models, the DSM and the NSM reveal almost the same potentials and energies, we conclude that both are adequate to describe the nonlocal effects. This is the reason why we reduce all further comparisons to the NVM and the DSM. We decided to use the DSM, because this model is implemented in both numerical methods, the EJIIM and the BEM.

7.2 Nonlocal electrostatic application to biomolecules

In this section, we apply the nonlocal theory on various biomolecules, namely the enzyme trypsin (PDB-entry 2PTC [88]), two CA variants of the HIV virus (homology models based on the N-

terminal domain of HIV-1 CA, PDB-entry 1GWP [73,128]), and the protein ovalbumin (PDB-entry 1OVA [120]). The molecular surfaces of trypsin and ovalbumin is exemplarily shown in Fig. 7.6.



(a) (b)

Fig. 7.6: Protein surfaces for (a) trypsin and (b) ovalbumin.

7.2.1 Trypsin

With its 3223 atoms, trypsin is a small protein. Trypsin is known for its functionality as a digestion enzyme and for its electrostatics to play a crucial role in the binding process of the inhibitor BPTI [77]. In Fig. 7.6(a) the binding pocket of trypsin can be located in the upper left.

7.2.1.1 EJIIM and BEM comparison

First, we solve the nonlocal electrostatic equations of the DSM with BEM and EJIIM to compare the influence of the boundary conditions and the overall accuracy of the solution. For the EJIIM calculation we took a mesh width of 0.5 Å and box dimensions [88Å× 88Å×88Å]. The boundary values have been calculated with the approximate boundary condition given in Definition 6.6 on p. 98. The BEM calculation is based on a surface triangulation with about 20000 elements, for which all the available memory was used.

The accuracy of both numerical methods is analyzed by means of the relative deviations of the BEM and EJIIM solutions for a randomly chosen point set as it was proposed in Section 6.4.2.2. In order to have representative point sets in Eq. (6.32) we stochastically choose 1000 points in Ω and Σ, respectively. Here, Σ is considered synonymous with the finite difference box [88Å×88Å×88Å]. For the point sets chosen inside the molecule, the relative difference $e(\phi_\Omega^{ejiim}, \phi_\Omega^{bem})$ equals 1.23e-2, whereas in Σ the difference has been $e(\phi_\Sigma^{ejiim}, \phi_\Sigma^{bem}) = 2.5\text{e-}3$. Therefore, we observe the *same* order of magnitude of the numerical error as for the molecules discussed in Section 6.4. From this, we can conclude that the EJIIM method can handle complex and non-trivial molecular geometries just as well as the BEM.

From the results in Section 6.4.2.4 we found that Eq. (6.19) is not only suitable for an approximation of the boundary values, but gives furthermore an impression of the electrostatic potential in the whole exterior region. In Fig. 7.7 we compare the numerical solution ϕ_Σ of the DSM and the NVM with the approximated potential ϕ_{approx}, i.e., evaluating Eq. (6.19) in Σ. As can be seen, the approximated boundary condition can reproduce ϕ_Σ of the DSM not only for simply shaped, small molecules, but, interestingly, it gives an adequate estimation of the potentials for complex

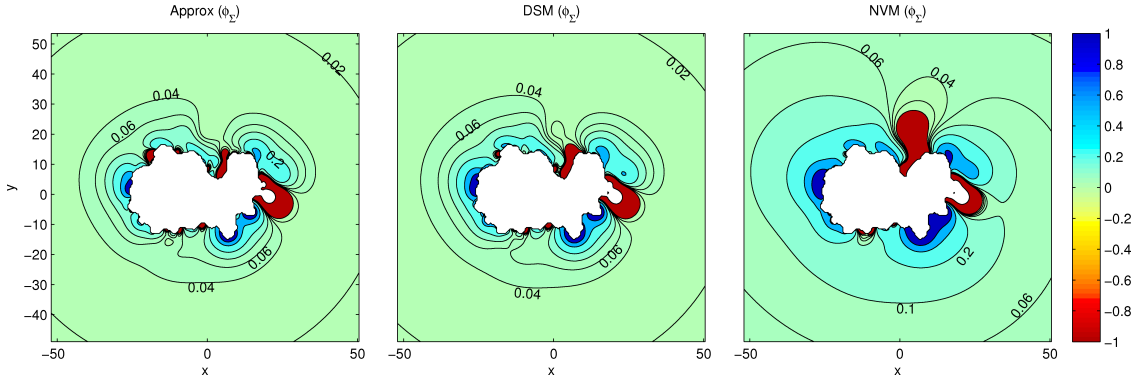


Fig. 7.7: The approximated boundary condition, Eq. (6.19) evaluated in Σ in comparison with ϕ_Σ of the DSM and ϕ_Σ of the NVM for an arbitrary cross section through the molecule.

geometries as well. This implies that we can reduce the box dimensions needed in the EJIIM calculations and therefore make the method more efficient. Moreover, it is a fast and very precise mean for a first estimation of the solution in the DSM, at least several Ångstrom apart from the molecule. As long as the outer potential is the focus of the study, as it is the case for the analysis of the ligand's trajectory around the biomolecule, this estimation of the electrostatic potential is an attractive alternative to the costly numerical calculations. As the underlying formulas are analytical expressions, this approximation can further be directly applied to biomolecules of arbitrary size without any requirements on numerical tools.

Although the boundary estimation was primarily designed for the DSM, we suggested in Remark 6.1 on p. 97 that Eq. (6.19) is a good approximation for all the other nonlocal models as well. However for trypsin, we see that although the correct order of magnitude for ϕ_Σ is achieved, the NVM results in higher potentials. This is a surprising result, because such a difference has not been observed for the small molecules, meaning that the reason must lie in non-trivial effects due to the complex shape of the molecule and its charge distribution.

7.2.1.2 Model comparison by contour surfaces

In contrast to the small molecules discussed in Section 7.1, proteins are composed of a high number of atoms that possess partial charges. This can cause a diverse electrostatic behavior: in Fig. 7.8, we depict ϕ_{mol} in a xy-plane through trypsin. In this plane, the binding pocket is found in the upper right. Although the molecular potential indicates the overall positive charge of trypsin, the binding pocket is locally characterized as a negatively charged region.

For this cross section, we now analyze ϕ^{reac} inside and ϕ_Σ outside of the protein calculated by the LM, the DSM, and the NVM. The calculations have been performed with the BEM, as this guarantees exactly the same input for all calculations. From the discussion before we suggest the NSM to yield roughly the same potentials as the DSM.

In Fig. 7.9 the reaction field potential ϕ^{reac} is plotted first. Although the charge pattern is much more complex for trypsin than for the test set molecules, the DSM once again predicts a homogeneous negative potential. Slight variations can be seen near the binding pocket and on the left side. These variations are stressed in the NVM demonstrating that the different boundary conditions of the two models clearly favor a different electrostatic behavior in Ω as expected in Remark 6.1. A behavior contrary to the NVM is given in the LM for the reaction field ϕ^{reac}: the negative potential, which is developed in the binding pocket turns into a decrease of the reaction field in the LM. Interestingly, this different behavior takes place near the binding pocket and

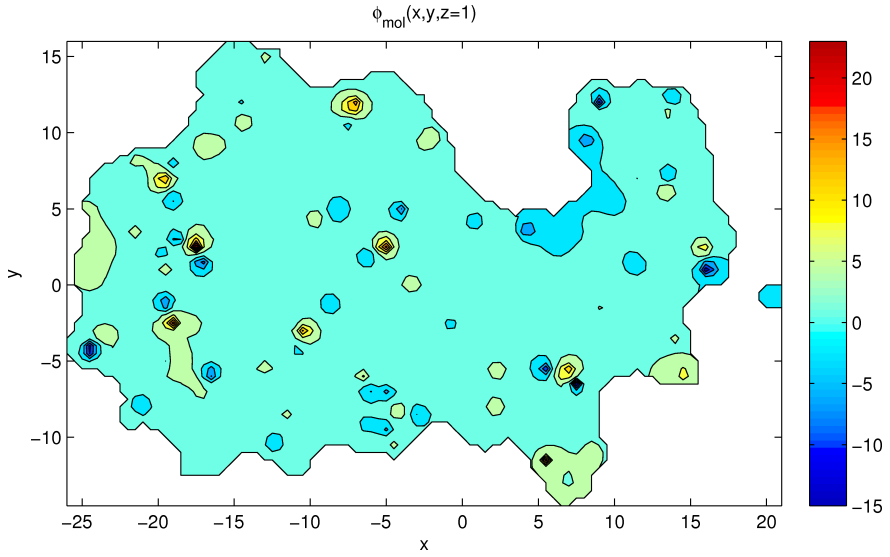


Fig. 7.8: ϕ_{mol} of a cross section through trypsin.

therefore we assume that the near-charge-effect together with the specially curved shape of the pocket are responsible for the difference.

Having in mind that the reaction field "takes care" of the transition condition between the electrostatic potential inside and outside the molecule, the reasons for the different behavior of ϕ^{reac} have to be likewise revealed in the outer potential shown in the lower part of Fig. 7.9: The binding pocket is shown in a way that the entrance area can be seen. For the DSM and in particular for the NVM a high, negative, electrostatic potential originates inside the binding pocket and

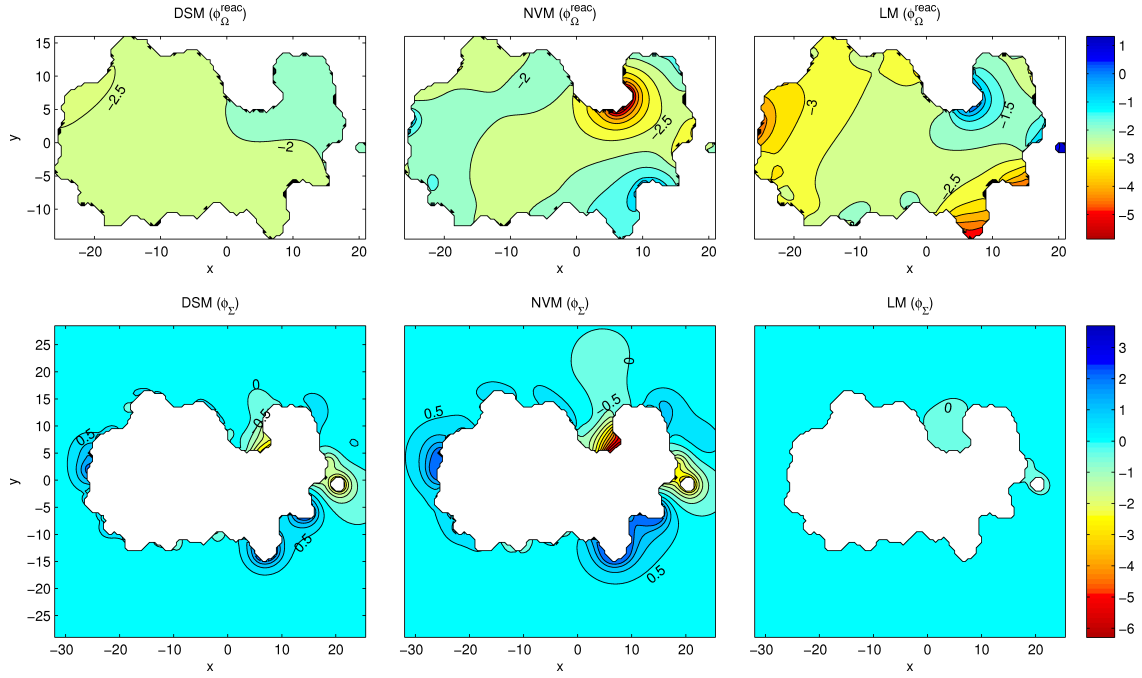


Fig. 7.9: Contour lines of a cross section through trypsin.

reaches further into Σ than in the LM. This fits to the reaction field ϕ^{reac} discussed before and moreover it gives us an interpretation: the binding pocket is a longish cylinder which points more than 10 Å inside the protein. Because of its curved shape, we can imagine that (a) the orientational degrees of freedom of the water molecules are broken or decreased as it is described by the vanishing Dirichlet boundary condition in the DSM and (b) the hydrogen bonds are primarily developed in normal direction out of the pocket, i.e., in the direction without disturbing dielectric boundaries.

7.2.1.3 Model comparison by isosurfaces

In Fig. 7.10 we plotted selected isosurfaces: for the NVM in (a) ±0.1V and (b) ±0.2V, for the DSM in (c) ±0.05V and (d) ±0.1V, for the LM in (e) ±0.05V and (f) ±0.1V.

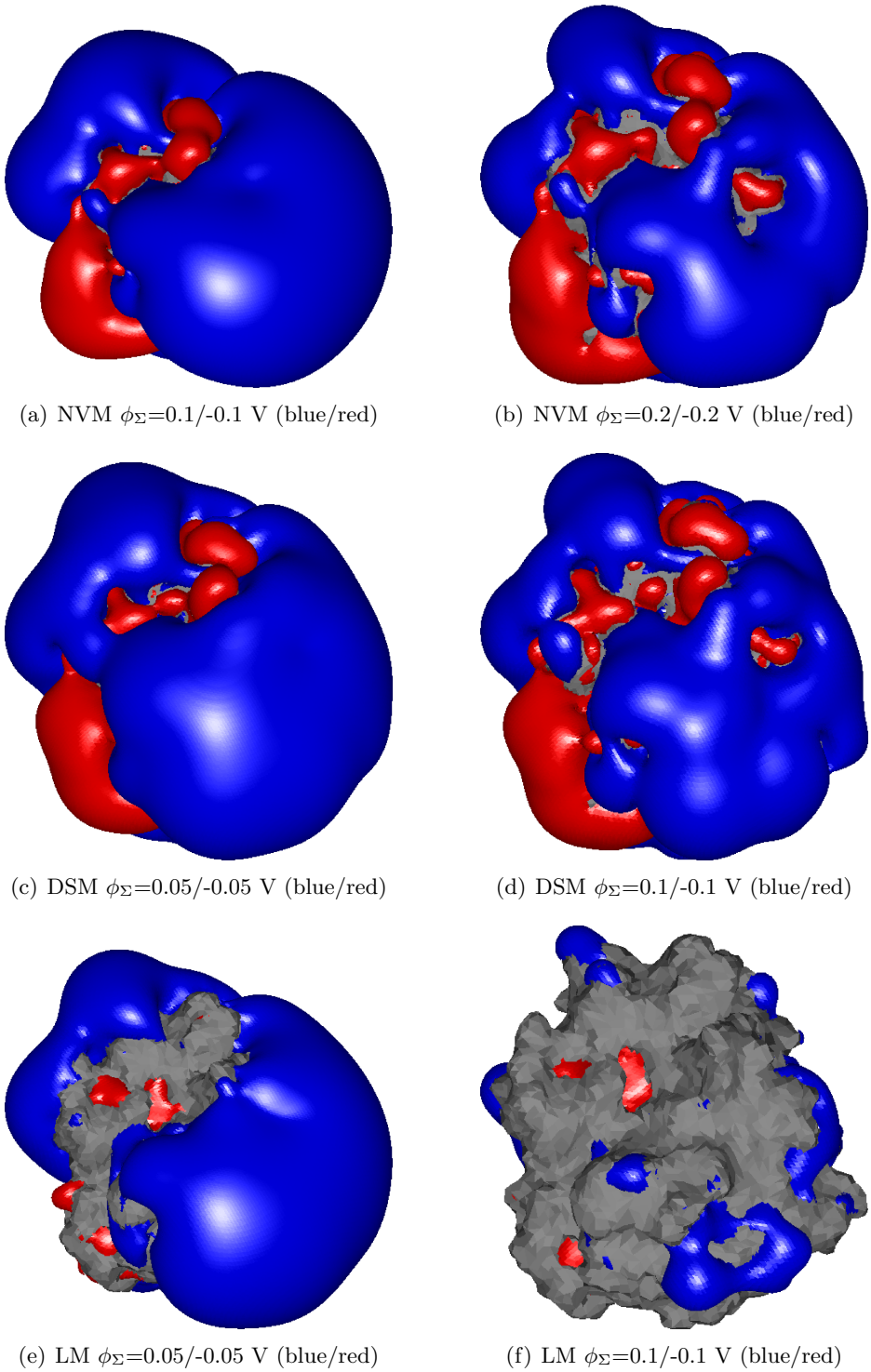
The view of the protein is the one shown in Fig. 7.6(a). The binding pocket for BPTI can be easily identified in (e) as the region with negative potential in the upper part.

First, we focus on Fig. 7.10(a,d,f), i.e., where we took the same isosurfaces for the three different models in order to compare the influence of the water network. The significant difference in the strength of the electric potential can once again be observed: for the LM (Fig. 7.10f), the potential reaches a value of -0.1V only deep inside the binding pocket. In the local theory, the water screens the electrostatic potential with a factor ∼78, but this is overestimated because the water molecules try to align with their next neighbors and this reduces its dielectric response. In both models, the NVM and the DSM, the negative potential is primarily developed inside the binding pocket and finds its way on a drawn-out path, which is flanked by regions of positive potential. Overall, both models exhibit the same long-range character.

However, as we have already found in the previous discussion, the NVM predicts an even stronger negative potential than the DSM. The isosurface -0.1V is more extended in space compared to the one in the DSM. In order to quantify the difference, we plotted in Fig. 7.10(b) the isosurface ±0.2V calculated by the NVM. Comparing these surfaces with those in Fig. 7.10(d) for the DSM, we see that the NVM predicts a potential roughly double in magnitude compared to that of the DSM (also confirmed by the comparison of Fig. 7.10(c) and (a)). Interestingly, this mainly holds for the positive isosurface, which originates from trypsin's total charge $Q = 6e$. Such an apparent difference has not been observed for the small molecules, and thus we assume that when the charge distribution and the shape of the molecule in consideration get more complex, the nonlocal models become more sensitive.

Further, we plotted the isosurfaces ±0.05V for the LM and the DSM. The reason is to demonstrate that for small electrostatic fields and small variations of the electrostatic potential, the local and the nonlocal models more and more agree with each other. This was one of the constraints for a suitable model of the dielectric nonlocal operator. Indeed, when comparing Fig. 7.10(c) and (e), we see that this is the case for the DSM.

In summary, the long-range potential and the characteristics established near the entrance area of the binding pocket, which are predicted by the nonlocal models, can have an important influence on the binding process of ligands: on the one hand, the ligand "feels" the force farther away. On the other hand, ligands can be bound more tightly to the binding pocket, since near the surface the nonlocal potential takes values up to one order of magnitude higher than the local potential.



(a) NVM ϕ_Σ=0.1/-0.1 V (blue/red)

(b) NVM ϕ_Σ=0.2/-0.2 V (blue/red)

(c) DSM ϕ_Σ=0.05/-0.05 V (blue/red)

(d) DSM ϕ_Σ=0.1/-0.1 V (blue/red)

(e) LM ϕ_Σ=0.05/-0.05 V (blue/red)

(f) LM ϕ_Σ=0.1/-0.1 V (blue/red)

Fig. 7.10: Isosurfaces of trypsin.

7.2.2 SIV/HIV capsid - primate TRIM5α interaction

The human immunodeficiency virus (HIV) is a virus that causes acquired immunodeficiency syndrome (AIDS), a condition in humans in which the immune system begins to fail, leading to

life-threatening opportunistic infections. One promising line of research is given by the recognition of *intrinsic immunity*. This means the anti-viral activity of proteins which are always present in species-specific living cells [17]. These intrinsic immune proteins or *restriction factors* constitute defense mechanisms of the host blocking the virus from replication and thus being potential agents against viral attack. As an example TRIM5α (Tripartite interaction motif five, splice variant α) is one of the most studied intrinsic immune proteins with respect to HIV and simian immunodeficiency virus (SIV), the analog of HIV for monkeys. TRIM5α recognizes the capsid proteins (CA) of entering viruses and prevents viral uncoating and reverse transcription by an up to now only rudimentarily understood mechanism [122]. The rhesus monkey TRIM5α variant is able to recognize and prevent HIV infection, whereas the human TRIM5α protein can prevent SIV infection [123]. This variation helps to explain why HIV and SIV infect humans and monkeys, respectively. The study of the differences in TRIM5α as well as HIV (CA) variants is a promising way to find more information on the detailed, specific interaction which causes a potential restriction or - vice versa - which improves the replicability of the virus [73]. For instance, according to experimental data, the SIVmac239 variant was not restricted, whereas the GH123 variant was strongly restricted by the primate TRIM5α. The research group of T. Shioda (Department of Viral Infection, Osaka University) found that mutations on the three loop regions of the CA variants are responsible for this different interaction behavior [73]. Besides a desired structural complement of the HIV binding site for the protein TRIM5α, an important factor for their interaction might be an electrostatic attraction of the protein to the specific binding site. This motivated to investigate the structural as well as the physical properties of the loop structures of SIVmac239 and GH123.

In a second application, we calculate nonlocal electrostatic potentials of two HIV-1 capsid variants, namely GH123 and SIVmac239. Besides the electrostatic calculations, we provided triangulations of the SAS, and also of the "SAS" with a probe radius of 3Å. The calculations were performed for K. Bozek (Max-Planck-Institute for Informatics, Computational Biology and Applied Algorithmics, Saarbruecken), who analysed the structural and the physical characteristics of SIV/HIV capsid - primate TRIM5α interactions in collaboration with the group of T. Shioda.

In Fig. 7.11 we illustrate both CA variants, GH123 (yellow) and SIVmac239 (red), where we

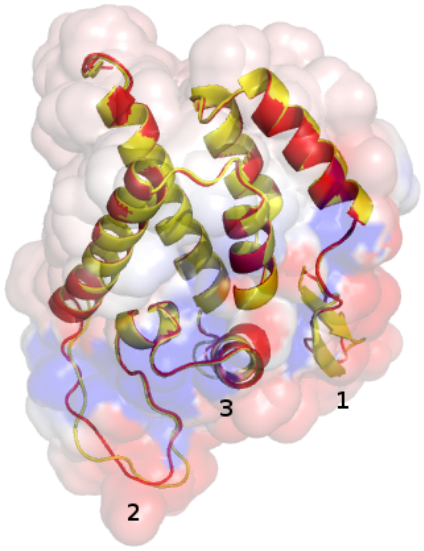


Fig. 7.11: Superposition of the GH123 CA (yellow) and the SIVmac239 CA (red) structures. The three loops interacting with TRIM5α are numbered. The SAS of both structures is transparently shown.

labeled the three loop structures in the upper part. The SIVmac239 has a more contracted shape compared to the expanded GH123 loop structure. A qualitative interpretation of the electrostatic differences can be given from inspection of Fig. 7.12, where the nonlocal electrostatic potentials on the SAS are shown for both variants. The loop regions are found in the lower part as it is noted in Fig. 7.11. Comparing the two variants, we see that in the loop regions 2 and 3 of GH123 a positive potential is predominant, whereas in SIVmac239 these regions exhibit a diffuse pattern. Qualitatively, this also holds for the local approach as well (not shown). However, in the nonlocal approach the potentials have a higher strength and the positive potential regions are much more separated from the negative potential regions. This can explain the *highly specific* interaction between the CA binding site and TRIM5α which sensitively hinges on point mutations of the building blocks of the loops.

7.2.3 Ovalbumin

As a third application of nonlocal electrostatics, we chose the protein ovalbumin, which is used in different areas of research such as immunization and biochemical studies. With its 6000 atoms, ovalbumin forms a complex surface shape, see Fig. 7.6(b). As a qualitative surface triangulation would have up to 50000 elements, such a calculation would exceed the memory requirements needed in our BEM implementations preventing its solution. However, with EJIIM we can solve system (6.2): for this, we use a grid resolution of 0.5 Å and box dimensions [93Å×93Å×93Å]. The boundary values have been calculated with the approximate boundary condition, see Section 6.2.4.1.

In Fig. 7.13 we show the isosurfaces of the electrostatic potential of the DSM for ±0.1V. Because of the large amount of partial charges near the surface, the potential exhibits a complex, diversified behavior. Altogether, the positive isosurface is closer to the surface. The molecule has a total charge $Q = -6e$, and this is why the negative contribution dominates the potential.

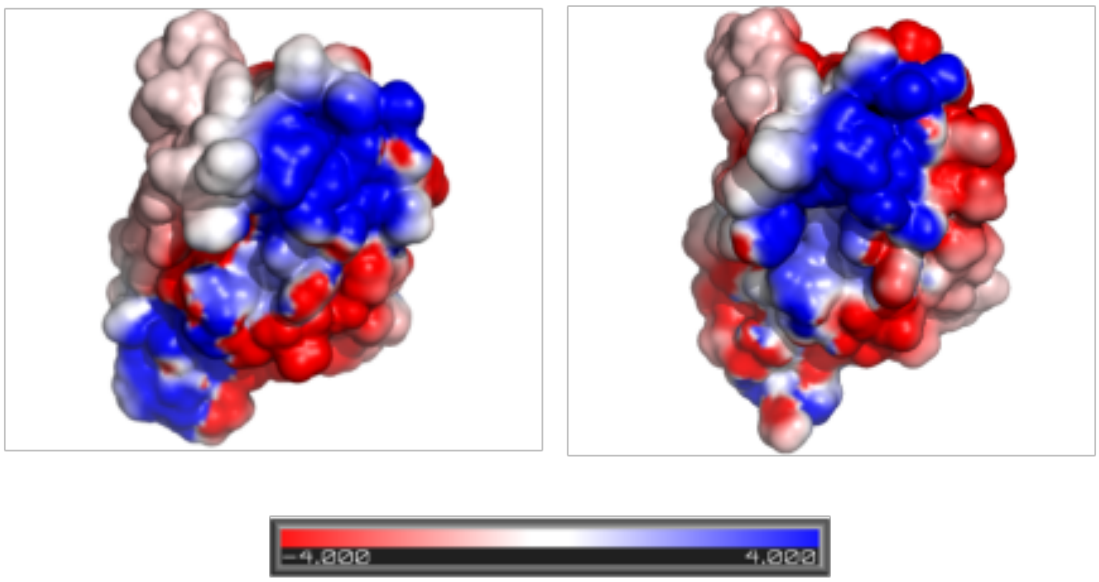


Fig. 7.12: The SAS of the CA variants GH123 (left) and SIVmac239 (right) colored by nonlocal electrostatic potential (DSM) in units of K_bT/e, the chosen color scale is shown below. Physical parameters: ($\varepsilon_\infty = 1.8\varepsilon_0$, $\varepsilon_\Omega = 10\varepsilon_0$, $\varepsilon_\Sigma = 80\varepsilon_0$, $\lambda = 3.028$Å).

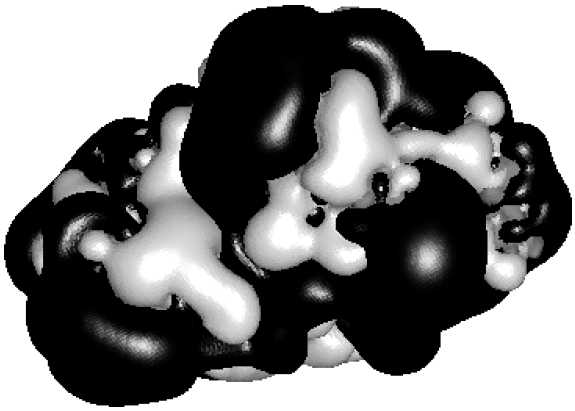


Fig. 7.13: Isosurfaces of ovalbumin for ϕ_Σ=0.1 V (white) and ϕ_Σ=-0.1 V (black).

The results underline that an application to biomolecules of arbitrary size and shape is possible with EJIIM. The input generator together with the EJIIM solver can be used to automatically generate the physical potentials. This is an essential step for further studies, especially for the comparison with experimental data.

7.3 Summary

In this chapter, we analyzed the nonlocal models which we derived before. The outcome of this study is that all the nonlocal models predict qualitatively and quantitatively the same, strong electrostatic potential around the considered molecules. Compared to the local model, the electrostatic potential is up to one order of magnitude higher, which is explained by the decrease in screening because of the water network in Σ.

First, we focused on the molecules of the test set. These have only a few partial charges and a simply shaped surface. The electrostatics predicted by the nonlocal models are comparable in terms of magnitude and shape. Slight variations in the reaction field ϕ_Ω^{reac} of the NVM have been explained by surface charges induced by molecular charges next to the dielectric boundary. This so called near-charge effect is caused by the boundary condition of the correlation field $\boldsymbol{F}$ on Γ in the NVM, i.e., the continuous extension of $\boldsymbol{F}_\Sigma$ into Ω. This condition admits variations of the water correlations that locally lower the outer dielectric response. In contrast, the reaction fields predicted by the two scalar models show an almost homogeneous behavior, which is caused by the fixed boundary condition of, for instance, a vanishing correlation field. When the charge pattern gets more complex as it is the case for FUDPOJ or trypsin, these aspects obviously play a non-trivial role.

Nevertheless, the analysis of the electrostatic contribution to the solvation free energy, which directly depends on the reaction field, reveals the same order of magnitude in all nonlocal models for all molecules of the test set. This clearly demonstrates that the variations due to the boundary conditions and thus their influence on physical quantities - at least on the electrostatic contribution to the solvation free energies - are small.

Second, we considered the small protein trypsin, where we focused on the electrostatics near the binding pocket. Once again, the nonlocal models predict an interesting long-range potential. Despite its overall positive total charge, several Ångstroms away from trypsin's binding pocket, this potential is negative and therefore favors the binding process of the positively charged BPTI. The strengthening of the negative potential in the entrance area is explained by the special characteris-

tics of the binding pocket: the strong negative charge distribution next to the dielectric boundary Γ and its cylindrical shape.

We additionally applied the boundary condition "without inner medium", which we deduced for the EJIIM in Section 6.2.4. This boundary condition yields approximations of the electrostatic potential in excellent agreement with the calculations of the DSM in Σ. We theoretically find this result for the case that the inner dielectric response ε_Ω is of the same order as the electronic response ε_∞ (see Remark 6.1). For $\varepsilon_\Omega \approx \varepsilon_\infty$, we thus conclude that the proposed boundary condition is a fast tool to estimate the outer nonlocal potential. It offers, for instance, the possibility to efficiently analyze the dynamics of a ligand which docks into the binding pocket of a protein. In this way, the approximated boundary condition takes the role of a generalized Born model for nonlocal electrostatics. Since it further does not require a numerical solver at all, it is possible to directly incorporate the analytical formulas as a first estimation of the nonlocal electrostatic potential into a biomolecular software package.

Furthermore, a cooperation with K. Bozek (MPI, Saarbruecken) gave us the possibility to perform electrostatic calculations to gain insight into the electrostatic differences of two HIV variants. We applied the EJIIM as well as the BEM, where for the latter we used surface triangulations generated by the algorithm discussed in Section 6.3.

As last application we chose the protein ovalbumin. This protein is still small, however, it has twice the volume of trypsin. Our BEM implementations of the DSM as well as of the NVM cannot cope with ovalbumin anymore, for they would require a triangulation which exceeds the memory capacity of our work station. The electrostatic nonlocal potential we showed is successfully calculated with the EJIIM.

From these applications, it is difficult to answer the question for an "optimal" nonlocal model: all the nonlocal models result in higher potentials near Γ and they reveal the local response when the potentials become small in Σ. The variations among the nonlocal models are small in contrast to the difference in one order of magnitude comparing the nonlocal to the local models. That is why we conclude that all of the nonlocal approaches are appropriate to incorporate the correlation effects of the hydrogen bond network. The method of choice therefore depends on the demand on efficiency and runtime. However, compared to the numerical complexity required to solve the vectorial model [39], these findings clearly support the use of a scalar model of nonlocal electrostatics.

The localized differences in magnitude of the potentials predicted in the various models originate from a different handling of the boundary effects. In fact, the NVM as well as the scalar nonlocal models have boundary conditions on Γ, which we motivated theoretically. For a further detailed modeling process of these boundary conditions, we need a detailed, experimental analysis of the boundary effects, which then can be incorporated in the numerical solver. In the NVM and the DVM two limits of the nonlocal behavior on the boundary are realized: the first neglects the disturbances on the surface and thus admits for variations of the nonlocal correlations, whereas in the second approach the network is frozen on Γ. The correct behavior lies between these approaches and depends on the locally differing polarity of the amino acids contributing to the surface.

The last two applications finally demonstrate that the newly developed numerical method as well as the algorithm proposed to generate the EJIIM input and the surface triangulations are efficient and easily applicable. Therefore, these algorithms are the basis for successful future work on nonlocal solvent studies of biomolecules.

Chapter 8

Electrostatics of biomolecules in ionic solution

Up to now, we considered the biomolecule to be immersed in pure water. However, in realistic conditions, i.e., in living organisms, there are also mobile salt ions present in the solvent. This chapter focuses on the question what happens electrostatically when a biomolecule is immersed in an aqueous solution containing ions.

In Section 8.1, we introduce the Poisson-Boltzmann equation to describe electrostatic effects of ionic solutions. We briefly recall the major approximations, which are required for the derivation of this *mean-field* theory.

The incorporation of the linearized Poisson-Boltzmann equation into the framework of (non)local electrostatics is done in Section 8.2: we extend the local model (Theorem 5.3.1) as well as the Dirichlet Scalar model of nonlocal electrostatics (Theorem 5.1.1) to account for ionic effects.

Despite water-water correlations, there might be ion-ion as well as ion-water correlations present in the biomolecular system. Section 8.3 addresses a single, very interesting effect, the so called *reentrant condensation of proteins*, which is one example for such non-trivial interactions causing a highly complex and sensitive behavior of proteins in ionic solutions. Based on experimental findings, we develop and apply a heuristic model that qualitatively describes the reentrant condensation of proteins. In this approach, though, we neglect the water-water correlations, and instead treat water as a high, constant dielectric. The rationale behind this neglect is that, in the situation described, the ion-protein correlation is expected to clearly dominate the overall behavior in the biomolecular system and thus has to be apparent for the commonly used water model as well.

8.1 Mean field theory of ionic solutions

Let us assume to have N point-like ions or particles with charge Ze present in the biomolecular system. In contrast to the biomolecule, which is described by a *fixed* charge density ρ, the positions of these mobile ions depend on the interactions among themselves as well as on the interactions with the biomolecule. This means that we have to solve a many-particle problem.

To account for the N-particle interactions, and, in this regard, to determine their probability distribution in the system, time consuming simulations are required. However, the reality is often well approximated by replacing the N-particle distribution function by a product of N identical one-particle distribution functions. The purpose of this *mean-field* approximation is to remove the correlations between the N particles and finally to find the ion density distribution $n(\boldsymbol{r})$ from minimizing the free energy functional [14, 56]. The minimization procedure yields

$$n(\boldsymbol{r}) = n_0 e^{-Ze\phi(\boldsymbol{r})/(k_{\mathrm{B}}T)}\,, \tag{8.1}$$

where n_0 denotes the bulk density of ions, ϕ the electrostatic potential of the biomolecular system, and $k_{\mathrm{B}}T$ the thermal energy. In Eq. (8.1) we consider the ions to be distributed in dependence of their potential relative to their thermal energy. Intuitively, this means that counter-ions, i.e., ions

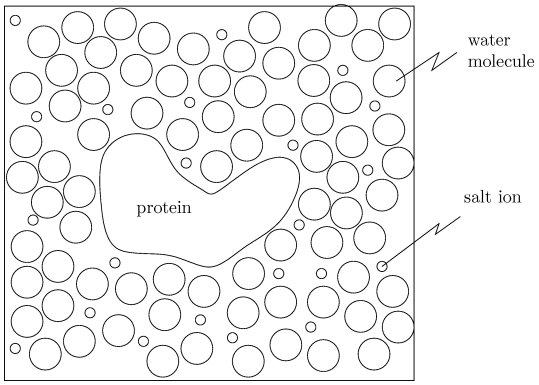


Fig. 8.1: In the native state, the protein is immersed in ionic water.

with a contrary charge with respect to the biomolecule, build up a layer near the surface of the biomolecule in order to minimize their potential energy. Due to the thermal energy of the ions and the resulting thermal motion, this layer is diffuse.

As Eq. (8.1) is an additional source of charge, the source term in Theorem 2.2.1 consisting of the molecular charge distribution ρ has to be extended: assuming different kinds of ions to be present in the solution results in the following material equation,

$$\boldsymbol{\nabla} \cdot \boldsymbol{D}(\boldsymbol{r}) = \rho(\boldsymbol{r}) + \sum_i Z_i e\, n_i(\boldsymbol{r}) \qquad \boldsymbol{r} \in \mathbb{R}^3 \tag{8.2}$$

with

$$n_i(\boldsymbol{r}) = n_{0,i} e^{-Z_i e \phi(\boldsymbol{r})/(k_{\mathrm{B}} T)} \qquad \boldsymbol{r} \in \Sigma\,, \tag{8.3}$$

where $n_{0,i}$ denotes the bulk density of ions of type i with valence Z_i.

Debye and Hückel discussed the diffuse double layer for spherical objects in contact with an electrolyte bath of monovalent co- and counter-ions [32]. They linearized the Poisson-Boltzmann equation assuming that the potential ϕ is not too strong anywhere in the system, i.e., the charges involved are small. The linearization of the ion density in Eq. (8.2) yields

$$\boldsymbol{\nabla} \cdot \boldsymbol{D}(\boldsymbol{r}) = \underbrace{\rho(\boldsymbol{r})}_{\text{molecular charge density}} + \underbrace{\left(-\chi_\Sigma\, \kappa_{sc}^2 \varepsilon_\Sigma \phi(\boldsymbol{r})\right)}_{\text{ion charge density}}, \qquad \boldsymbol{r} \in \mathbb{R}^3 \tag{8.4}$$

with

$$\kappa_{sc} := \sqrt{\frac{e^2}{\varepsilon_\Sigma k_{\mathrm{B}} T} \sum_i n_{0,i}\, Z_i^2} \tag{8.5}$$

$$\chi_\Sigma := \begin{cases} 0 & \text{in } \Omega \\ 1 & \text{in } \Sigma\,. \end{cases} \tag{8.6}$$

In Eq. (8.6) we introduced the characteristic function χ_Σ to restrict the ion distribution to Σ. Further, with the definition of κ_{sc} in Eq. (8.5), a new length scale, the so called Debye screening length,

$$\lambda_{sc} = 1/\kappa_{sc}\,,$$

is introduced into the system. Its physical meaning is revealed in Section 8.2.1 and Section 8.2.2, where we state the analytical formula of the potential ϕ of the Born model in ionic solvents.

A typical concentration $n_0 = 0.15 \frac{mol}{l}$ of monovalent salt immersed in bulk water, which is characterized by its macroscopic dielectric constant, $\varepsilon_\Sigma = 78.5\varepsilon_0$, yields a Debye screening length

of $\lambda_{sc} = 7.8$Å. This corresponds to $\kappa_{sc} = 0.12\frac{1}{\text{Å}}$.

In summary, the Poisson-Boltzmann model is based on the following assumptions [56]:

- The solvent degrees of freedom are represented by the continuum dielectric response and a specific ion-water molecule correlation is assumed to be negligible.
- Ions are effectively modeled as point charges.
- Ion-ion interactions in the solvent are assumed to be negligible.

In spite of the limitations in application and interpretation which result from the simplicity of the mean field approach, Poisson-Boltzmann models are widely and successfully used for numerical calculations of the screened, electrostatic potential of molecules in ionic solutions [2, 9, 70]. They provide an deeper understanding of the electrostatic ionic screening in biomolecular systems. For low charged, dilute systems the Poisson-Boltzmann theory yields correct results [56] and even the condensation of ions on a charged surface is partly captured in these theories [99].

8.2 Extension of the electrostatic models to ionic solutions

In the previous section, we learned that, under certain conditions, the ionic solvents are well described by an additional charge density,

$$\rho_{ion}(\boldsymbol{r}) = -\kappa_{sc}^2 \phi(\boldsymbol{r}) \,, \quad \boldsymbol{r} \in \Sigma \,, \tag{8.7}$$

where we used the definition of κ_{sc} given in Eq. (8.5). This means that the incorporation of ionic effects into the local as well as into the nonlocal models of electrostatics consists of the following change

$$\rho(\boldsymbol{r}) \;=\; \left\{ \begin{array}{ll} \rho_{mol}(\boldsymbol{r}) & \boldsymbol{r} \in \Omega \\ 0 & \boldsymbol{r} \in \Sigma \end{array} \right\} \quad \rightarrow \quad \rho(\boldsymbol{r}) \;=\; \left\{ \begin{array}{ll} \rho_{mol}(\boldsymbol{r}) & \boldsymbol{r} \in \Omega \\ \rho_{ion}(\boldsymbol{r}) & \boldsymbol{r} \in \Sigma \end{array} \right\} . \tag{8.8}$$

Since the basic material equations given in Theorem 2.2.1 on p. 18 do not change by the extension of the charge distribution proposed in Eq. (8.8), we can directly state the following theorem for local electrostatics with ionic screening.

Theorem 8.2.1 *Model for local electrostatics with ionic screening: Extending the charge density as proposed in Eq.* (8.8) *in the local model (LM) of electrostatics (Theorem 5.3.1) yields*

$$\left.\begin{array}{rcll} \varepsilon_\Omega \Delta \phi_\Omega &=& -\rho_{mol} & \text{in } \Omega \\ \Delta \phi_\Sigma &=& \kappa_{sc}^2 \phi_\Sigma & \text{in } \Sigma \\ & & & \\ \varepsilon_\Omega \partial_n \phi_\Omega &=& \varepsilon_\Sigma \partial_n \phi_\Sigma & \text{on } \Gamma \\ \phi_\Omega &=& \phi_\Sigma & \text{on } \Gamma \,. \end{array}\right\} \tag{8.9}$$

In the following, the local model with linear Poisson-Boltzmann theory is abbreviated as LM_LPB.

In the second equation of system (8.9), the following Yukawa operator appears

$$\mathcal{L}_{\kappa_{sc}} := (\triangle - \kappa_{sc}^2) \,.$$

Here, the Yukawa operator $\mathcal{L}_{\kappa_{sc}}$ describes the weakening of the electrostatic potential by ions. This can be clearly seen on the electrostatic potential of the Born sphere with charge q and radius a:

$$\phi_\Omega = \frac{q}{4\pi}\left[\frac{1}{\varepsilon_\Omega\, r} + \left(\frac{1}{\varepsilon_\Sigma(\kappa_{sc}a+1)} - \frac{1}{\varepsilon_\Omega}\right)\frac{1}{a}\right] \qquad r < a \tag{8.10a}$$

$$\phi_\Sigma = \frac{q}{4\pi r}\left[\frac{1}{(\kappa_{sc}\, a+1)\,\varepsilon_\Sigma}\, e^{-\kappa_{sc}(r-a)}\right] \qquad r > a\,. \tag{8.10b}$$

The $\frac{1}{r}$-potential of the local model transforms into a Yukawa-like potential, which means that the ion density screens the Laplace-like potential. The screening is characterized by an exponential decay on the length scale of the Debye screening length λ_{sc} defined in Eq. (8.5).

The appearance of the Yukawa operator in the theory of ionic solvents is interesting, as in the nonlocal theory of water, a Yukawa operator appears as well. This reveals another physical meaning of the Yukawa operator in the screening theory of ions: it emphasizes that the distribution of the counter- and co-ions effects the electrostatic potential in a nonlocal way.

8.2.1 Dirichlet Scalar Model with ionic screening

In the last section, we introduced the LM_LPB, which is the commonly used electrostatic model to account for constant dielectric response and an additional ionic screening effect. A coupling between the nonlocal water correlations and the ionic screening might result in an interesting competition of these two effects. In order to study the interplay, we now extend the Dirichlet Scalar model to consider the additional source term given in Eq. (8.8).

Theorem 8.2.2 *Dirichlet Scalar Model with ionic screening: Extending the charge density as proposed in Eq.* (8.8) *in the Dirichlet Scalar Model (DSM, Theorem 5.1.1) of nonlocal electrostatics reveals*

$$\left.\begin{aligned}
\varepsilon_\Omega \triangle\phi_\Omega &= -\rho_{mol} && \text{in } \Omega\\
\triangle(\varepsilon_\infty\,\phi_\Sigma + (\varepsilon_\Sigma - \varepsilon_\infty)\,F_\Sigma) &= \varepsilon_\Sigma\,\kappa_{sc}^2\phi_\Sigma && \text{in } \Sigma\\
\phi_\Omega &= \phi_\Sigma && \text{on } \Gamma\\
\varepsilon_\Omega\partial_n\phi_\Omega &= \varepsilon_\infty\partial_n\phi_\Sigma && \text{on } \Gamma\\
\\
\mathcal{L}_\kappa F_\Sigma &= -\kappa^2\phi_\Sigma && \text{in } \Sigma\\
\partial_n F_\Sigma &= 0 && \text{on } \Gamma
\end{aligned}\right\} \tag{8.11}$$

In the following, this model is called the Dirichlet Scalar model with linear Poisson-Boltzmann theory. It is abbreviated as DSM_LPB.

In this section, we discuss in detail the DSM_LPB for the Born sphere (Section 8.2.1.1). Thereafter in Section 8.2.1.2, we develop an EJIIM solver for the numerical solution process and present first applications of the DSM_LPB to BRMW1 and FUDPOJ of the test set.

8.2.1.1 Born sphere

The solution of system (8.11) for the Born sphere of radius a with charge q is given by

$$\phi_\Omega = \frac{q}{4\pi}\left[\frac{1}{\varepsilon_\Omega\, r} + \frac{1}{a}((C_1 + C_2) - \frac{1}{\varepsilon_\Omega})\right], \qquad r < a \tag{8.12a}$$

$$\phi_\Sigma = \frac{q}{4\pi}\,\frac{1}{r}\Big[\underbrace{C_1\, e^{-p_-\,(r-a)}}_{\text{first branch}} + \underbrace{C_2\, e^{-p_+\,(r-a)}}_{\text{second branch}}\Big], \qquad r > a\,, \tag{8.12b}$$

where we used the following notations:

$$s := \frac{\varepsilon_\Sigma}{\varepsilon_\infty}(\kappa^2 + \kappa_{sc}^2) \tag{8.13a}$$

$$p_\pm^2 = \frac{1}{2}\left(s \pm \sqrt{s^2 - 4\frac{\varepsilon_\Sigma}{\varepsilon_\infty}\kappa^2\kappa_{sc}^2}\right) \tag{8.13b}$$

$$C_1 = \frac{\varepsilon_\Sigma\kappa_{sc}^2 - \varepsilon_\infty p_+^2}{(p_-^2 - p_+^2)(p_- a + 1)}\frac{p_-^2}{\kappa_{sc}^2\varepsilon_\Sigma\varepsilon_\infty} \tag{8.13c}$$

$$C_2 = -\frac{\varepsilon_\Sigma\kappa_{sc}^2 - \varepsilon_\infty p_-^2}{(p_-^2 - p_+^2)(p_+ a + 1)}\frac{p_+^2}{\kappa_{sc}^2\varepsilon_\Sigma\varepsilon_\infty} \tag{8.13d}$$

From the analytical solution of the spherically symmetric case, we see how the two effects, the ionic screening and the water-water correlations, are competing in weakening/strengthening the electrostatic potential: the electrostatic potential consists of two branches, which are specified by $\{C_1, p_-\}$ and $\{C_2, p_+\}$ in Eq. (8.12). Both parts have the shape of Yukawa potentials, $\sim \frac{e^{-p\,r}}{r}$. Recalling the outer nonlocal electrostatic potential of the Born solution in the DSM (see Section 5.3),

$$4\pi\,\phi_\Sigma = \frac{1}{\varepsilon_\Sigma r} + \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_\Sigma}\right)\frac{e^{-(r-a)/\lambda}}{r}\frac{1}{a/\lambda + 1}\,, \quad r > a\,,$$

we can directly interpret the change of the Laplace-like part to a Yukawa-like part, as a consequence of the additional ionic screening.

Fig. 8.2 shows the prefactors given in Eqs.(8.13)(b-d) for a Born sphere of radius $a = 1$Å with the standard parameters for the nonlocal response ($\varepsilon_\Sigma = 78.5\varepsilon_0, \varepsilon_\infty = 1.8\varepsilon_0$). We reduced the κ-κ_{sc} plane to the region of realistic values, i.e., $\kappa \sim 0.05\frac{1}{\text{Å}}$ and $\kappa_{sc} \sim 0.1\frac{1}{\text{Å}}$.

The functional shape of the prefactors is mainly directed by the limiting values, which we gain from the functional expressions (Eqs. (8.13)(b-d)):

$$\left.\begin{aligned} \lim_{\kappa\to\infty} C_1\, e^{-p_-^2\,(r-a)} &= \tfrac{1}{a\,\kappa_{sc}+1}\,\tfrac{1}{\varepsilon_\Sigma}\,e^{-(r-a)\kappa_{sc}} \\ \lim_{\kappa\to\infty} C_2\, e^{-p_+^2\,(r-a)} &= 0 \end{aligned}\right\} \quad\Leftrightarrow\quad \text{LM\_LPB}(\varepsilon_\Sigma) \tag{8.14a}$$

$$\left.\begin{aligned} \lim_{\kappa_{sc}\to 0} C_1\, e^{-p_-^2\,(r-a)} &= \tfrac{1}{\varepsilon_\Sigma} \\ \lim_{\kappa_{sc}\to 0} C_2\, e^{-p_+^2\,(r-a)} &= \tfrac{1}{a\kappa\sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}+1}\left(\tfrac{1}{\varepsilon_\infty} - \tfrac{1}{\varepsilon_\Sigma}\right)e^{-(r-a)\kappa\sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}} \end{aligned}\right\} \quad\Leftrightarrow\quad \text{DSM}(\varepsilon_\Sigma) \tag{8.14b}$$

$$\left.\begin{aligned} \lim_{\kappa\to 0} C_1\, e^{-p_-^2\,(r-a)} &= 0 \\ \lim_{\kappa\to 0} C_2\, e^{-p_+^2\,(r-a)} &= \tfrac{1}{a\kappa_{sc}\sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}+1}\,\tfrac{1}{\varepsilon_\infty}\,e^{-(r-a)\kappa_{sc}\sqrt{\varepsilon_\Sigma/\varepsilon_\infty}} \end{aligned}\right\} \quad\Leftrightarrow\quad \text{LM\_LPB}(\varepsilon_\infty) \tag{8.14c}$$

As we noted in Eqs. (8.14), the limits correctly merge with either the DSM for $\kappa_{sc} \to 0$ and the LM with macroscopic response ε_Σ for $\kappa \to \infty$, respectively. The limit $\kappa \to 0$ coincides with a maximization of the correlation length and this in turn freezes the system to the local, electronic response ε_∞.

At a glance, the complex behavior of the DSM_LPB for the Born sphere is governed by these limiting values. In the whole region, the nonlocal model is strongly dominated by the ionic screening: in both quantities, C_1 and C_2, we see the $\frac{1}{\kappa_{sc}}$-decrease and the linear increase of the exponents

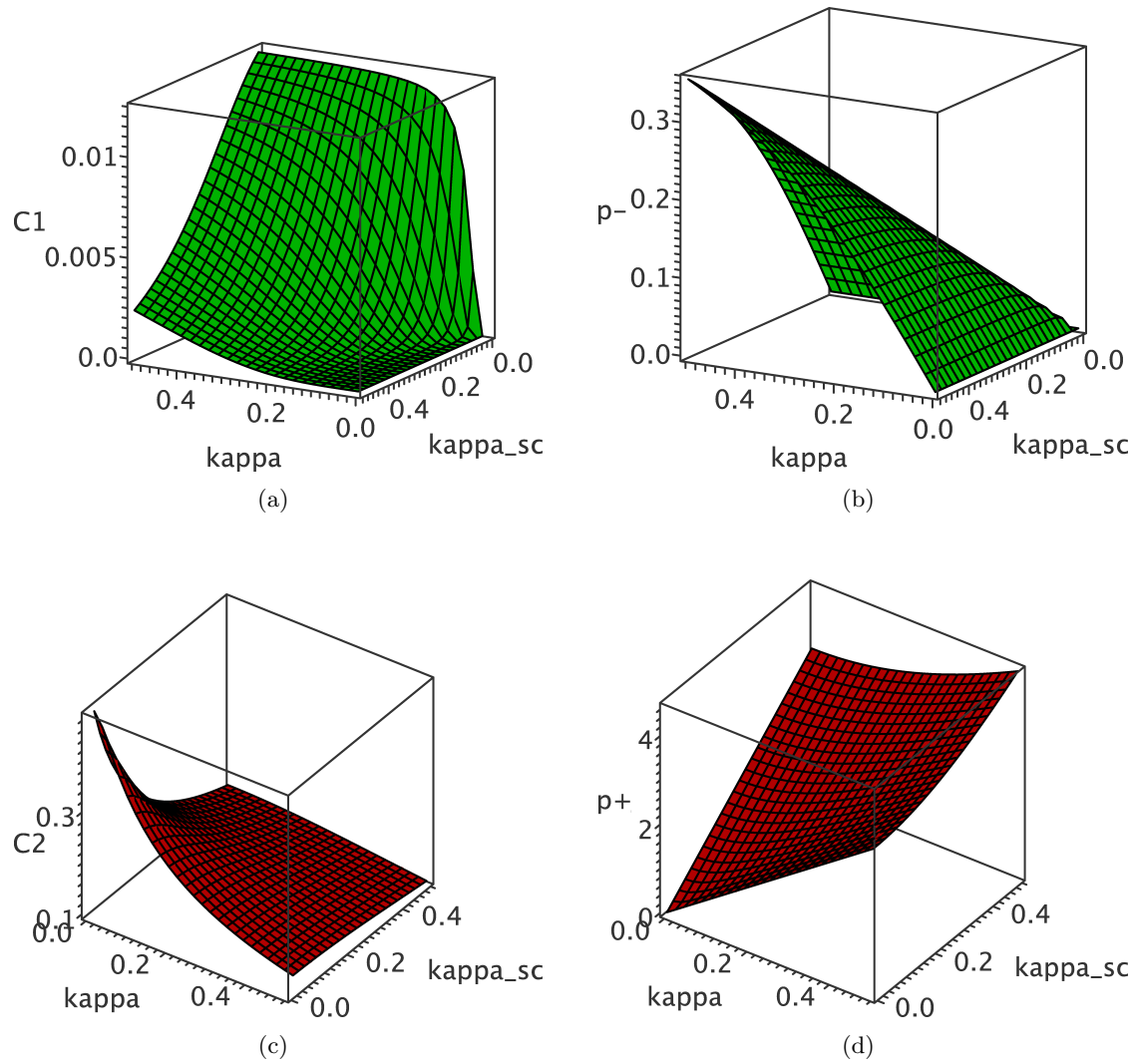


Fig. 8.2: Prefactors of the nonlocal/screening branches for the Born solution of the DSM_LPB. κ and κ_{sc} are plotted in units of $\frac{1}{\text{Å}}$.

$p_\pm$ for increasing κ_{sc} as it is implied by Eqs. (8.14). The prevalence of the ionic screening can be explained by the fact that on average the water correlations cause higher field values in Σ. This in turn increases the ion distribution around the molecule and therefore the ionic screening.

The linear increase of the exponent p_- in dependence of the ionic strength is blocked above a critical κ, which can be clearly seen on the exhibited nose in Fig. 8.2(b). A further increase of κ_{sc} lets p_- remain unchanged. This demonstrates the interplay of both effects in a non-linear way. From the DSM, we know that the first branch characterized by $\{C_1, p_-\}$ corresponds to a local, macroscopic contribution, which is regained for increasing λ (decreasing κ).

In the DSM, the exponent p_+ shown in Fig. 8.2(d) determines the nonlocal correlations and therefore it is clear that for increasing κ (decreasing λ), this contribution vanishes. The same is true for an increase in the ionic strength, as then there are more ions immersed in the solvent, effectively weakening the electrostatic potential. Both effects can be read off the linear dependence of p_+ on κ and κ_{sc}.

There is a remarkably strong decrease of C_1 for increasing ionic strength and small κ values. In

order to analyze its origin we consider a further limiting process in Eq. (8.14b) and Eq. (8.14c):

$$\begin{aligned}
\text{(8.14b)}: \quad \lim_{\kappa\to 0}\left(\tfrac{1}{\varepsilon_\Sigma r}+\tfrac{1}{a/\lambda+1}\,(\tfrac{1}{\varepsilon_\infty}-\tfrac{1}{\varepsilon_\Sigma})e^{-(r-a)/\lambda}\right) &= \tfrac{1}{r}(\tfrac{1}{\varepsilon_\Sigma}+[\tfrac{1}{\varepsilon_\infty}-\tfrac{1}{\varepsilon_\Sigma}]) = \tfrac{1}{\varepsilon_\infty r}\\
\text{(8.14c)}: \quad \lim_{\kappa_{sc}\to 0}\left((\tfrac{1}{a\kappa_{sc}\sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}+1})e^{-(r-a)\kappa_{sc}\sqrt{\frac{\varepsilon_\Sigma}{\varepsilon_\infty}}}\right) &= \tfrac{1}{r}(0+\tfrac{1}{\varepsilon_\infty}) = \tfrac{1}{\varepsilon_\infty r}
\end{aligned}$$

As we can see in the preceding equations, both limiting processes result in the same model, the local model with a pure electronic response ε_∞. However, the path to reach this limit is totally different and this in turn causes a discontinuity of C_1 and C_2. The DSM always consists of a Laplace-like and a Yukawa-like part, where the former corresponds to the LM and the latter to the nonlocal corrections. In contrast, the electrostatic potential of the LM_LPB consists of a single Yukawa-like part, which smoothly turns into the Laplace-like potential for vanishing ionic screening. The discontinuity of the prefactors C_1 and C_2 emphasizes that the interplay of the water correlations and the ionic screening is sensitive for small κ and κ_{sc}. When κ_{sc} gets larger, the overall behavior of the system is dominated by the ionic screening.

8.2.1.2 EJIIM model extension

In the last section, we deduced the DSM_LPB, i.e., the Dirichlet Scalar model with linear ionic screening effect (see Theorem 8.2.2) and analyzed its solution for the Born model. It reveals a strong screening effect and therefore a weakened electrostatic potential. In order to apply this model to arbitrarily shaped molecules, we have to set up a numerical solver.

In contrast to the BEM, where a modification or an extension of the nonlocal equations always requires a fundamental solution of the differential system, such a restriction is not given for the EJIIM. Moreover, the modifications we have to make in the DSM implementation to account for the linear ionic screening effect are marginal, because only the following field equation of the DSM given in Theorem 5.1.1 on p. 73 is affected:

$$\begin{aligned}
&\triangle(\varepsilon_\infty\,\phi_\Sigma+(\varepsilon_\Sigma-\varepsilon_\infty)\,F_\Sigma) = 0 && \text{in } \Sigma\\
\text{with ionic screening:}\quad &\triangle(\varepsilon_\infty\,\phi_\Sigma+(\varepsilon_\Sigma-\varepsilon_\infty)\,F_\Sigma) = \varepsilon_\Sigma\,\kappa_{sc}^2\phi_\Sigma && \text{in } \Sigma
\end{aligned} \tag{8.15}$$

The equations in Ω as well as the transition conditions remain as given in Theorem 5.1.1. This means that the involved part of the EJIIM implementation, i.e., the incorporation of the discontinuities on the dielectric boundary Γ, is already correct for the DSM_LPB. Only the discretization of the field equations has to be changed in Σ in agreement with Eq. (8.15).

Fig. 8.3 shows the outer electrostatic potential of the DSM_LPB applied to the small, charged molecule BRMW1 (upper row) and to the small, uncharged molecule FUDPOJ (lower row). We compare the DSM_LPB results with the corresponding results of the DSM, the LM, and the LM_LPB. As for both molecules the results are qualitatively the same, but more distinct for the charged molecule, we concentrate on BRMW1. Comparing the DSM with the DSM_LPB, we see that when introducing the ionic screening additionally to the water correlations, the electrostatic potential is considerably weakened by one order of magnitude a few Ångstrom away from the molecule. Compared to the local potentials seen on the right, the ionic screening is stronger as it was already anticipated when discussing the spherically symmetric system. The reason lies in the higher nonlocal potentials and the resulting high contribution of screening counter-ions on the molecular surface.

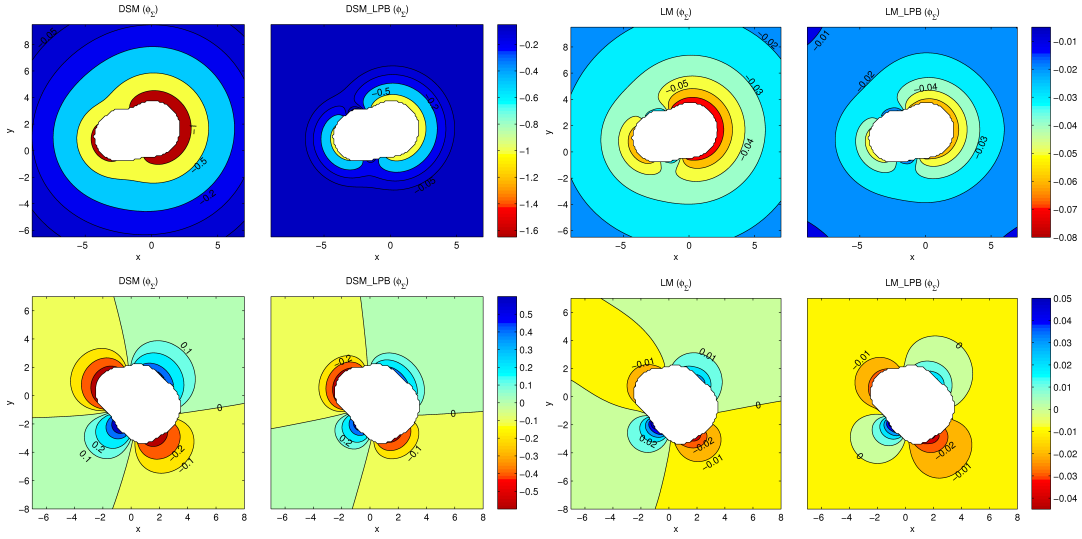


Fig. 8.3: Application of the DSM_LPB on BRMW1 (upper row) and FUDPOJ (lower row). Physical parameters ($\varepsilon_\Sigma = 78.5\varepsilon_0, \varepsilon_\infty = 1.8\varepsilon_0$, $\kappa = 0.05\frac{1}{\text{Å}}$, $\kappa_{sc} = 0.12\frac{1}{\text{Å}}$)

8.2.2 Conclusion

Within the last sections, we demonstrated the easy extensibility of the electrostatic model equations using the EJIIM: we focused on the more realistic molecular system which additionally contains solvated ions. The study of the analytically solvable, spherically symmetric system gave an estimation of the interplay of ionic screening and water correlation effects. Further, it served for testing the changes made in the EJIIM implementation in order to account for the ionic screening. In summary, we found that the ions have a considerable influence on the nonlocal electrostatic potential of small molecules. This qualitative study has to be seen as a starting point of further investigations, which among others have to consider the following two important aspects:

1. The correlation length λ of water and the ionic screening length $\frac{1}{\kappa}$ are taken to be independent of each other in the previous studies. A more realistic assumption, however, is that the mobile ions non-trivially influence the correlation length of the water molecules. When approaching the molecular surface, the counter-ion density can increase in a way that, on average, several ions are present within the correlation sphere of the water molecules [1] and thus the free correlation length of water can decrease.

2. At sufficiently high ionic densities, steric effects prevent ions from accumulating at charged interfaces to the extent predicted by the standard PB theory which we introduced before [14, 60]. Steric constraints lead to saturation of the ion density near the interface and, thus, increase their concentration in the rest of the interfacial region. A first approach to account for the steric effect is to introduced the so called *Stern* or *ion exclusion layer*, which is the layer in the solvent where the ions cannot further approach the molecular surface because of their finite size. In common Poisson-Boltzmann finite difference solvers, such finite size effects are accounted for by redefining the grid which stores the ionic screening factor κ_{sc} to have non-vanishing values only outside the Stern layer. As the EJIIM *is* a finite difference solver, it is possible to incorporate the ion exclusion effect into the existing EJIIM solvers.

[1] See, for example, the estimations of the ion density on a planar charged surface in [60], Section 12.4, p. 218.

8.3 Beyond mean field theory: multivalent ions in protein solutions

In this section, we further concentrate on ionic solutions, but as the focus will lie on the ions, we now treat the dielectric effects of water as in the LM_LPB (see Theorem 8.2.1), which means that we model the solvent as a homogeneous, constant dielectric medium. In the last section, we discussed the effects of counter- and co-ions in the framework of the linearized Poisson-Boltzmann equation and we learned that statistically distributed ions have an important impact on the range of the electrostatic field.

In Section 8.2.2 the introduction of a Stern layer was suggested as a possibility to account for finite size effects, which lead to a modified ion distribution near dielectric and charged interfaces. However, this is only a small concession to the complexity of the counter- and co-ions: ion-ion, ion-water, and specific ion-molecule correlations are still missing in the mean field theory and this neglect leads to the fact that many interesting effects of ionic as well as of acidic and basic solutions cannot be explained. For instance, the variance in pH can cause proteins to change their charge state, or influence the stability of their secondary structure [111]. Another prominent example is the functionality of metal ions in solution: they are responsible for nucleophilic catalysis in enzymes, for the electron transfer in proteins, for the stabilization of protein structure, and for the stimulation and blocking of excitable cells [5, 57, 95].

In this section, we focus on the highly interesting phenomenon of reentrant condensation of protein solutions, which is caused by the non-trivial interactions of multivalent ions and therefore requires beyond-mean-field theories for its explanation. Reentrant condensation generally describes the phenomenon that biomolecules in solution undergo a condensation upon adding a critical concentration of a so called condensation agent - typically counter-ions of the biomolecules. The condensed phase redissolves for a second, higher critical concentration of the condensation agent. Fig. 8.4 depicts such a phase transition as a function of the counter-ion and the molecule concentration. The critical counter-ion concentrations separating the different regions are defined as c^* and c^{**}.

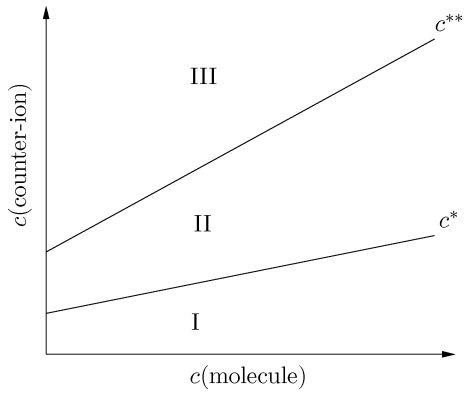


Fig. 8.4: The phase diagram: I) solved phase, II) condensed phase, III) dissolved phase.

Reentrant condensation has already been observed in systems like DNA or polyelectrolytes with multivalent ions such as the naturally occurring polyamines spermidine^{3+} and spermine^{4+} and the inorganic cation $Co(NH_3)_6^{3+}$ [18]. Here, the phenomenon of the condensation and a re-dissolution of the biomolecules originates from ion-ion correlations, i.e., correlated fluctuations in the ion atmosphere around the macroion [46, 56, 103].

Recently, reentrant condensation has also been found in protein solutions [150], where the chemical agents which cause the condensation effect, are trivalent rare earth metal ions, such as La^{3+} and Y^{3+}. A deeper understanding of the attractive protein-protein interaction going along with

reentrant condensation of proteins could be pathbreaking for some human diseases where the agglomeration of protein plays a crucial role, for instance, Alzheimer disease. This makes the reentrant condensation of proteins a new and important field of research.

At first view, the high valence of the counter-ions supports that reentrant condensation is driven by ion-ion correlations as it is the case for DNA and polyelectrolytes. However, the interaction between DNA and La^{3+} has been measured [126] and it turns out that for very small amounts of La^{3+}, a specific discrete binding of the trivalent metal ions to the PO_2^- groups on the backbone of the DNA exists. The binding causes the DNA to condense but not to redissolve in solution. This observation indicates that (a) the metal ions interact in a different - more specific - way than typical condensation agents of DNA and (b) the theories for reentrant condensation, which are established for DNA, are *not* directly transferable to protein solutions.

A reason could lie in fact that DNA and protein strongly differ in their shape and in their charge distribution: the first is well described as a rod with a constant negative surface charge whereas the latter is characterized by a complex protein-specific shape and a diverse charge distribution. The structural difference can be clearly seen in the sketch of the E1-DBD dimer bound to a DNA double helix (PDB-entry 1KSY) in Fig. 8.5.

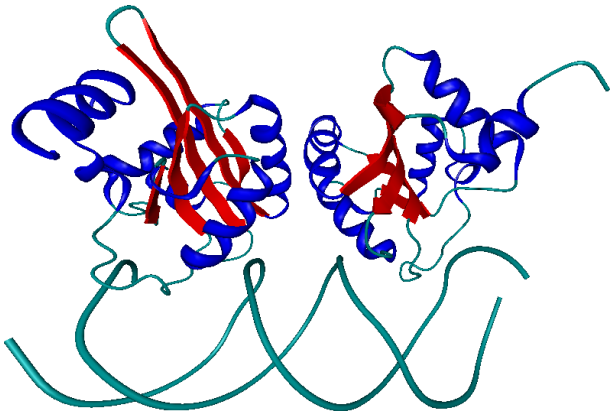


Fig. 8.5: The structural difference of protein (upper) and DNA (lower).

The difference of DNA and protein reentrant condensation and further the novelty of the latter motivates first to carefully study experimental findings of protein solutions with multivalent metal ions and thereafter to develop a model which is based on this knowledge. Thus, before we derive a simple, heuristic model in Section 8.3.2 to describe the transition of a protein solution into the condensed phase, we start in Section 8.3.1 with a summary on experimental results, which give rise to the assumption that multivalent Y^{3+} ions have a high, specific affinity to negatively charged amino acid side chains. Based on this assumption, we implement a modified titration simulation, which calculates the probability of the multivalent ions to be bound to the surface of the molecule. In Section 8.3.3, we apply this titration model to the protein bovine serum albumin (BSA) and further discuss the results for a set of four distinct proteins in reference to experimental measurements in Section 8.3.4. The comparison confirms that a specific binding of counter-ions on the protein's surface reveals the main features of reentrant condensation.

8.3.1 Experimental findings

The affinity of metal ions to the functional groups of molecules can be found in many biochemical processes and therefore its study is a widely spread and important field of research. Examples are the complexation of metal ions with a set of surrounding molecules or anions [26, 41, 115] or the

adsorption of metal ions on surface exposed amino acid side chains, which is used for example to purify protein solutions by the Immobilized Metal Ion Affinity Chromatography [131, 136].

8.3.1.1 High binding affinity of multivalent metal ions

In particular, the specific binding of the rare earth metal ions such as Y^{3+} and La^{3+} to carboxyl residues has been investigated in pharmaceutical and biochemical studies. The following observations have been made:

- The presence of a small amount of multivalent Yttrium ions Y^{3+} inhibits the calcium channel in excitable cells. The blocking stems from the high affinity of Yttrium to the calcium binding site, which is modeled in literature as a site composed of two negatively charged ions at a distance of 2Å or specified as an arrangement of two Glutamic acids [65, 95]. In [65], the authors state that
  "*the Yttrium cation has a similar ionic radius to calcium, and thus exhibits a good geometric match to the selectivity filter, but is trivalent and has a higher binding coefficient for carboxyl residues than calcium, likely leading to its very high potency as a channel inhibitor*".

- The specific binding of Y^{3+} to proteins has been studied in titration experiments [23,81,145]: in these studies a strong affinity of the Yttrium ions to surface exposed, negatively charged side chains has been observed. The stability constant of Y^{3+} to Tyrosine, for instance, ranges between $pK_m = 5.09$ ($I = 0M$, $35°C$) and $pK_m = 4.43$ ($I = 0.1M$, $25°C$) [108].

8.3.1.2 Reentrant condensation of proteins

The previously mentioned experimental studies give reason to believe that individual multivalent metal ions exhibit a specific correlation with surface exposed, negatively charged amino acids of proteins. In the following, we recapitulate the experimental findings of the recently described reentrant condensation of proteins [150, 151] and give an interpretation with respect to an existing ion-amino acid correlation effect:

1. **Experiment**: For a set of four proteins - all with a negative total charge - the reentrant behavior is found when the multivalent counter-ion concentration is consecutively increased.

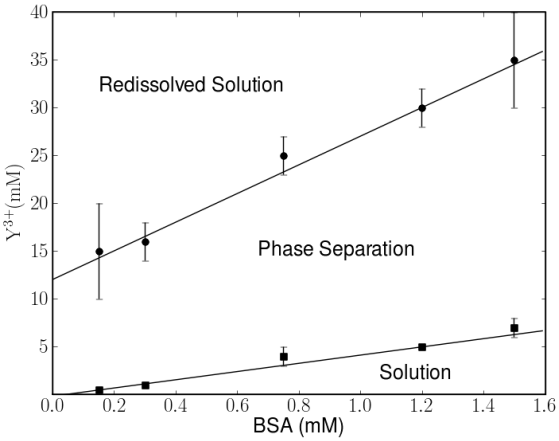


Fig. 8.6: The phase diagram for BSA and Yttrium ions.

In order to describe the counter-ion concentration where the protein enters and leaves the condensed phase, respectively, we introduced the functions c^* and c^{**} and refer to them in the following as *phase transition curves*. In the experiments, both phase transitions, c^* and c^{**}, conform well with a linear relation to the protein concentration c_P [2]. Therefore, we make the ansatz

$$c^*(c_P) = c_1 + m^* c_P \tag{8.16}$$
$$c^{**}(c_P) = c_2 + m^{**} c_P \,, \tag{8.17}$$

where c_1, c_2, m^* and m^{**} are constants. In Fig. 8.6, such a phase transition is shown for BSA. The reentrant condensation is induced by the counter-ions Y^{3+} of the salt YCl_3. The presented data is retrieved from experiment [151], and listed in Tab. 8.1 for a set of four analyzed proteins.

Tab. 8.1: Experimental results of the phase transitions for the protein set [151].

	experimental results			
Protein	m^*	m^{**}	c_1 [mM]	c_2 [mM]
BSA	4.30±0.5	15±1	-0.2 ±0.2	12 ±1
HSA	3.40±0.5	58±2	0.1 ±0.2	3 ±1
OVA	1.80±0.2	4±1	0.1 ±0.2	24 ±5
BLG	0.50±0.2	3±1	0.05±0.02	0.08±0.02

Interpretation: It is worth to note that up to c^*, the concentrations of protein and salt differ at most by the factor m^*, i.e., the number of Yttrium ions and protein molecules in solution is almost equal. This supports the assumption to treat the counter-ions as ligands for the localized binding sites. Further, the linearity of the first phase transition curve can be interpreted as a near-quantitative binding of the Yttrium ions to the available protein binding sites in solution.

2. **Experiment**: By electrophoretic mobility experiments, the zeta potentials for the set of proteins in Tab. 8.1 have been determined as a function of salt concentration. For all analyzed proteins, the zeta potentials exhibit a zero-crossing close to the Y^{3+} concentration where the protein enters the condensed phase (see Fig. 8.11 and [151]).

   Interpretation: For colloids, one often uses the zeta potential as a measure of the effective surface charge in ionic systems. It indicates the movement of the ion-macromolecule complex in an external electric field. In this way, the zero-crossing of the zeta potential near the first phase transition curve can be interpreted as an effective charge inversion of the protein-ion agglomerates.

3. **Experiment**: A first structure determination of the protein crystal of β-Lactoglobin (BLG) with Y^{3+} reveals that in the condensed phase (a) the globular structure of the protein remains stable and (b) a small number of Yttrium ions are localized in the neighborhood of three negatively charged amino acid side chains [149].

   Interpretation: The study on the crystal structure gives explicit evidence that a specific interaction of the multivalent counter-ions and the negatively charged residues on the protein

[2] Note that for c^{**}, this does not hold when the salt changes the pH of the solution as it is the case for $AlCl_3$ [151].

surface is involved in the reentrant condensation of proteins.

The experimental results of reentrant condensation discussed above and the treatment of multivalent ions as selective binding partners of negatively charged side chains in former studies support that, at least up to the first phase transition, the reentrant condensation phenomenon is driven by the specific binding of multivalent counter-ions on the protein's surface: the binding of the positively charged counter-ions changes the charge distribution of the protein - counter-ion agglomerate. In particular, the protein experiences a charge inversion for a critical metal ion concentration. Within the picture of counter-ion binding, the transition to the condensed phase is induced by attractive forces, which do not become relevant until the repulsive electrostatic forces become small. An increase of the counter-ion concentration above the first phase transition causes the electrostatic repulsions once again to increase and this in turn causes the condensed phase to finally dissolve.

This simple explanation motivates a theoretical model for the counter-ion binding. The model, which we developed, is explained in detail in the following. It is based on the introduction of an additional energy contribution, the energy upon binding of metal ions to negatively charged, surface exposed side chains. This energy contribution is described by a stability constant K_m similar to the dissociation constants K_a of amino acids in protonation studies.

8.3.2 Theoretical model

Analogous to the titration used in the experiments to determine the stability constant of Y^{3+} to an isolated amino acid side chain [108], we propose a generalized titration model simulating the specific binding of counter-ions to amino acid side chains in the protein environment.

The theoretical background of specific interactions at surface exposed amino acids is well established for the protonation of proteins in basic or acidic environment, where the protonation state is calculated in titration simulations [3, 12, 111]. The aim of a titration simulation is to find the optimal proton binding configuration for varying pH, i.e., to calculate the proton binding state of minimal free energy. The dissociation constant of a titratable amino acid side chain i in the protein is denoted $K_{a,i}$. The determination of $K_{a,i}$ is the key to find the optimal binding state. The details on the titration formalism can be found in the original articles on K_a evaluation [11, 111]. Additionally, we give an introduction of titration theory in Appendix 10.3.

In our study, we want to determine the optimal metal ion binding configuration as a function of metal ion concentration. This binding takes place at unprotonated (and therefore negatively charged) aspartic and glutamic acids lying on the surface of the protein. Their specific dissociation constant is denoted $K_{m,i}$ for the i^{th} possible binding site. We therefore define in Tab. 8.2 the nomenclature for the theoretical description of multivalent metal ion binding in analogy to the nomenclature used in titration simulations. Substituting the quantities given in the left column in Tab. 8.2 by the quantities in the right column corresponds to switching from protonation to metal-

Tab. 8.2: Nomenclature transferred from the titration community; pK_x is defined as the negative logarithm of the dissciation constant of acid X.

protonation		ion binding	
proton	(H^+)	metal ion	(Y^{3+})
amino acid	(e.g., given in [111])	amino acid	(Asp, Glu)
K_a		K_m	
pH		pM	

ion binding. In this way, the formalism we now briefly derive for the metal-ion binding results in the same formulas for the free energy difference.

8.3.2.1 Electrostatic free energy of the binding state

The Hamiltonian commonly used to account for the free energy in titration studies is given by the electrostatic contributions only. The electrostatic potentials are typically obtained via a finite difference Poisson-Boltzmann scheme (FDPB) [3, 12, 111], where additionally immersed salt ions are treated in a mean-field approximation neglecting ion-ion correlation effects.

Compared to titration studies, the situation of binding multivalent ions to the protein's surface exposed amino acids is energetically more complicated, since besides the protein-ion correlations, ion-ion interactions between the strongly charged ions are also present in the system. These interactions contribute to the total free energy and influence the overall behavior of the solutes [103]. However, as stated in Section 8.3.1, we have reason to believe that for the low salt concentrations realized in the experiments, the ion-protein affinity is much greater than the ion-ion correlations. Thus, we model the free energy of the system in the way it is done in protonation simulations. Of course, the simulation protocol which we will develop is also applicable when additional energy contributions are taken into account.

We combine the linear Poisson-Boltzmann theory with specific interactions between the strongly positive ions and negatively charged sites of the protein to incorporate the correlation effect of multivalent ions and amino acid side chains. Taking this approach, the metal ions play two different roles: on the one hand, they can bind to surface exposed amino acids. This is realized by the introduction of an energy contribution which favors the ion-binding to special side chains. On the other hand, the counter-ions as well as the co-ions of the added salt slightly screen the electrostatic field of the protein when solving the linear Poisson-Boltzmann equation.

As introduced previously, the residue i releases the metal ion with a specific dissociation constant, $K_{m,i}$, which results from the interplay between various forces such as entropic, hydrophobic, van der Waals, and electrostatic forces. Since these forces sensitively depend on the ion distribution and the chemical environment of the protein and further on the strength of the protein-metal ion complex, it is hardly possible to determine $K_{m,i}$ in experiment. Analogous to protonation theory, we calculate $K_{m,i}$ as a function of the "electrostatic protein environment" in which the side chain i is embedded [3, 12, 111]. To this end, we follow the titration protocol described in detail in [111] and in Appendix 10.3. Before the protocol is now adapted to the binding of metal ions, we have to introduce a few definitions:

Definition 8.1 Active residue or active site:
An *active residue* denotes a side chain of the protein that can bind a metal ion. A priori, we assume that an active residue is negatively charged and therefore unprotonated. To account for the binding process of a metal ion with charge $+m\,e$ and radius r, the active residue i can occupy one of two possible states:

$$s_i = \begin{cases} 0: & \text{the unbound form, which is the residue in the unprotonated state. Additionally, we add a sphere with radius } r \text{ and charge 0 to the active residue.} \\ 1: & \text{the bound form, which is the residue in the unprotonated state. Additionally, we add a sphere with radius } r \text{ and charge } +m \text{ to the active residue.} \end{cases} \tag{8.18}$$

As defined in Eq. (8.18), we keep the model as simple as possible and do not consider volumetric or structural changes after the binding of the ion, but let the binding process only affect the charge state. In the unbound state, the charge of the active site is $-e$, whereas in the bound state it is $(m-1)\,e$.

To account for the free energy of solvation, we introduce the so called *reference state*, on which the change in energy is based. This reference state is given by the protein in its protonation state of initial pH and vanishing metal ion concentration, i.e., the state of all active sites being in their unbound state.

Definition 8.2 The standard dissociation constant pK_m^{stnd}:
Analogous to the definition of pK_a^{stnd}, which describes the release of a proton from a certain standard peptide, we denote by pK_m^{stnd} the negative logarithm of the experimentally measured dissociation constant of a metal ion with respect to the standard peptide. The dissociation constant K_m^{stnd} can be expressed by the difference of the free energies of the bound (1) to the unbound state (0):

$$K_m^{stnd} = e^{-\Delta E_{stnd}^{diss}(1\to 0)/k_B\,T}$$

or expressed by the binding energy gained when changing from the unbound to the bound state:

$$K_m^{stnd} = e^{+\Delta E_{stnd}^{bind}(0\to 1)/k_B\,T} \tag{8.19}$$

The standard peptide represents the amino acid in a neutral environment without the protein background. For an amino acid X, this is, for instance, the N-acetyl N-methylamide derivative, *ACE-X-NME*, where *ACE* and *NME* denote the neutral acetyl and methyl end caps, respectively [12, 147]. In our analysis we used the value of p$K_m^{stnd} = 5.09$ for all active sides. This value has been extrapolated from experimental data for the binding of Y^{3+} to Tyrosine [108].

Definition 8.3 The intrinsic dissociation constant pK_m^{intr}:
The p$K_{m,i}^{intr}$ is defined as the negative logarithm of a hypothetical dissociation constant for the protein, assuming all active sites are kept in their reference state except for site i. Substituting $K_{m,i}^{stnd}$ by $K_{m,i}^{intr}$ in Eq. (8.19) yields the binding energy in the "intrinsic state" of the protein.

With these definitions, we energetically formulate the binding process: assume a protein with N active sites and let $\boldsymbol{s}$ be a vector that assigns to every active side chain i the state s_i (the reference state is given by $\boldsymbol{s} = \mathbf{0}$). Then, the p$M$-dependent free energy difference, $\Delta G_{solv}(\mathbf{0} \to \boldsymbol{s}, \mathrm{p}M)$ between the protein in charge state $\boldsymbol{s}$ and the reference state is given by

$$\begin{aligned}\Delta G_{solv}(\mathbf{0}\to\boldsymbol{s}, \mathrm{p}M) = (\ln 10)\, k_B\, T \sum_{i=0}^{N} s_i\,(\mathrm{p}M - \mathrm{p}K_{m,i}^{intr}) \\ + \sum_{i=0}^{N}\sum_{j>i}^{N} s_i\, s_j\, W_{ij},\end{aligned} \tag{8.20}$$

with

$$W_{ij} \;:=\; E_{ij}(1,1) - E_{ij}(1,0) - E_{ij}(0,1) + E_{ij}(0,0)\,.$$

Here, $E_{ij}(s_i, s_j)$ denotes the interaction energy of the side chain i with side chain j

$$E_{ij}(s_i, s_j) = \sum_l q_{(l,s_j)}\phi_{(i,s_i)}(\boldsymbol{r}_l)\,,$$

where $\boldsymbol{r}_l$ is the position of atom l in the active site j. The electrostatic potential $\phi_{(i,s_i)}$ originates from the charge distribution of the active residue i in charge state s_i, where all the other charges of the protein are set to zero.

As $(\mathrm{p}M - \mathrm{p}K_{m,i}^{intr})$ measures the energy of binding a metal ion at site i when all other active sites are in their reference state, the first part of Eq. (8.20) has to be corrected by the second part, adding those interaction energies that correspond to the current state vector $\boldsymbol{s}$. Since p$K_{m,i}^{intr}$

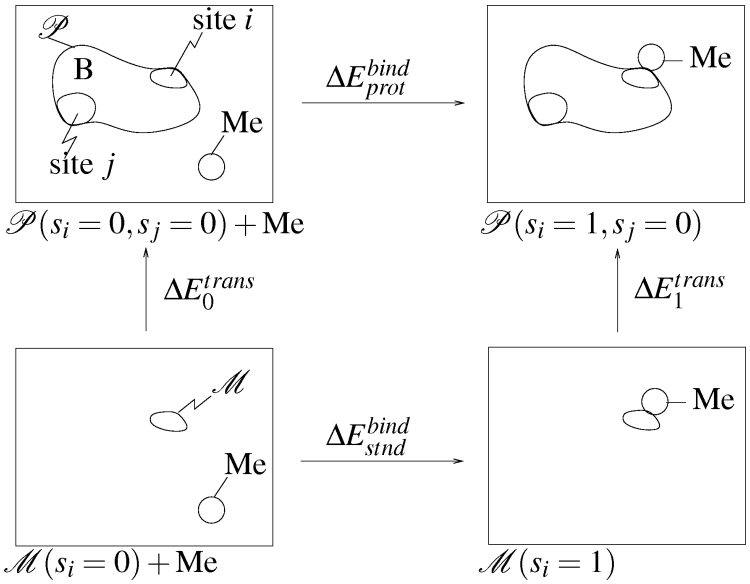


Fig. 8.7: Thermodynamic cycle to calculate $\mathrm{p}K^{intr}_{m,i}$ from $\mathrm{p}K^{stnd}_{m,i}$.

depends on the chemical environment of the protein, it differs from the experimentally measurable $\mathrm{p}K^{stnd}_{m,i}$ of site i. By considering the thermodynamic cycle illustrated in Fig. 8.7, $\mathrm{p}K^{intr}_{m,i}$ can be calculated from $\mathrm{p}K^{stnd}_{m,i}$: the idea is to "move" the active residue i from the standard peptide into the protein environment ($\mathcal{P}$), which comprises other active sites j and the background amino acids (B). Further, the model compound $\mathcal{M}$ represents the standard peptide of the active amino acid i. In our study, the isolated amino acid i defines the model compound as it is usually done in protonation simulations [111]. The transfer energies to bring the active residue into the protein environment depend on whether a metal ion (Me) is bound or not. This is expressed by ΔE^{trans}_1 and ΔE^{trans}_0, respectively. We thus have

$$-(\ln 10) k_B T(\mathrm{p}K^{intr}_{m,i} - \mathrm{p}K^{stnd}_{m,i}) = \Delta E^{trans}_1 - \Delta E^{trans}_0 \,. \tag{8.21}$$

With Eq. (8.20) and Eq. (8.21), we describe the change in free energy when changing the state vector $\boldsymbol{s}$, where we describe the free energy only by the electrostatic contributions. It is mainly governed by two characteristic quantities: the vector $\boldsymbol{pK}^{intr}_m$, comprising the $\mathrm{p}K^{intr}_{m,i}$ of every active residue i, and the interaction matrix W_{ij}.

8.3.2.2 Monte Carlo simulation of the specific metal ion binding

The aim of the numerical simulation is the calculation of quantities like the average effective charge of the protein-ion complex, $\langle Q_{\mathrm{eff}}(\mathrm{p}M)\rangle$ or the probability of finding the site i in the binding state s_i, $\langle s_i(\mathrm{p}M)\rangle$ for a given ion concentration. In statistical mechanics, the partition function encompasses all this information of the system [111]

$$Z(\mathrm{p}M) = \sum_{\boldsymbol{s}} \exp\left(-\frac{1}{k_B T}\, G(\boldsymbol{s}, \mathrm{p}M)\right) . \tag{8.22}$$

From Eq. (8.22), we can compute the relevant ensemble averages, for instance, the probability distribution of active site i to be occupied

$$\langle s_i(\mathrm{p}M)\rangle = \frac{1}{Z(\mathrm{p}M)} \sum_{\boldsymbol{s}} s_i \exp\left(-\frac{1}{k_B\,T}\,G(\boldsymbol{s},\mathrm{p}M)\right), \tag{8.23}$$

where a substitution $G(\boldsymbol{s},\mathrm{p}M)$ to $\Delta G_{solv}(\mathbf{0}\rightarrow\boldsymbol{s},\mathrm{p}M)$ in Eq. (8.22) and Eq. (8.23) does not change $\langle s_i(\mathrm{p}M)\rangle$.

If we assume N active sites, the partition function consists of 2^N contributions. When N exceeds about 30 sites, it becomes impractical to exactly calculate the partition function. Since N already exceeds this threshold for small proteins, we sample the state space with a standard Monte Carlo simulation [16, 111]. Different from what is proposed in [16], we neglect the effect of strongly interacting residues, since the size of the metal ions prevents a simultaneous switching.

In order to formulate the Monte Carlo procedure, we define the *MC step* as the result of randomly changing N_{dof} times one degree of freedom in the state vector, where N_{dof} is the number of all degrees of freedom of the system - in our case, the number of active sites. This procedure allows to assume that the results of two consecutive MC steps are independent from each other. The Monte Carlo algorithm proceeds for every discrete pM as follows:

initialization: the simulation starts with an arbitrary state vector $\boldsymbol{s}$.

equilibration: we execute N_{equi} MC steps, where, due to the Metropolis criterion, the new state vector is rejected or accepted [16]. The acceptance/rejection of an attempt to change the binding state of a residue is based on the free energy difference given in Eq. (8.20); in case of rejection, the old state vector serves as input for the next run, else the new one.

data collection: we take the result of the equilibration as initial input and execute N_{coll} MC steps. Based on the Metropolis criterion, we reject or accept the result of the current MC step as before. Additionally, every accepted step is part of the data we store for postprocessing – the calculation of the probability distribution of binding.

8.3.3 Implementation

In this section, we focus on the general preprocessing of the proteins and the numerical methods we use for the calculations. Our numerical investigation is adapted to experimental studies, where the effect of increasing YCl_3 concentration and therefore the effect of Yttrium ions on a set of four protein solutions is analyzed [150, 151]. Here, the BSA-Y^{3+} interaction serves as the reference example.

8.3.3.1 Initial protonation state

The structure of the protein in question is retrieved from the Protein Data Bank (PDB, [15]) and checked for completeness. Further, we remove crystallographic water molecules and alternate side chain locations.

As for BSA, no crystal structure is available, we constructed a homology model of BSA based on the known structure of human serum albumin (HSA). This approach is justified by the large sequence identity of the two proteins (57%) that implies a shared fold. Homology modeling was performed using the software suite PRIME (Schrödinger Inc., New York, release 1.6.307).

In order to simulate the binding process, we have to know the initial charge state of the protein. In the beginning of the experiments, no salt is immersed in the neutral solution. Thus, the initial

charge state of the protein equals the protonation state at pH=7 and an ionic strength of zero. We predict the protonation state of all exposed acidic and basic side chains via the H++ webserver [2,45]. Charge distributions for the protein are taken from the AMBER force field and the positions of the hydrogens are optimized by H++.

Here, we use the inner dielectric response, ε_{in}, in a range from $2\varepsilon_0$ to $6\varepsilon_0$ in order to find a good agreement between the experimentally measured total charge of the protein and the predicted total charge of the H++ run. In the literature, we find a total charge of $-10\,e$ for BSA at neutral pH [110, 127]. With an inner dielectric response $\varepsilon_{in}=2\varepsilon_0$, the theoretically predicted charge state has a total charge of $-11\,e$, which is in good agreement.

8.3.3.2 Electrostatic calculations and Monte Carlo simulation

In order to efficiently calculate the electrostatic potentials, we decided to take the optimized software package APBS in the automatically-configured sequential focusing mode [9]. The energy of the protein is calculated in a box with 6 Å additional space to the bounding box of the protein. As boundary condition of the coarse calculations we use "Single Debye-Hückel".

The Monte Carlo method is written in `C++` and uses the BALL library for an easy handling and processing of the proteins [70]. The equilibration has been done with $N_{equi} = 40000$, and the data collection with $N_{coll} = 50000$. To give a rough impression on the computational costs, executing N_{equi} and N_{coll} MC steps for a molecule with 32 degrees of freedom, the calculation takes 0.94 seconds on a machine with four Intel(R) Xeon(R) CPU W3540 @ 2.93GHz processors and 12GB RAM.

8.3.3.3 Determination of active residues for metal ion binding

As we stated in Section 8.3.2, the set of active sites in our numerical simulation is composed of all surface exposed unprotonated Asp and Glu residues. The basis for this assumption has been given in Section 8.3.1. In both, the bound and the unbound state[3], we place the Yttrium ion symmetrically between the oxygens with a distance of 2.43 Å in accordance with the literature [80, 94], such that it is located in the plane spanned by the oxygens and the connected carbon. The Yttrium radius is taken to be 2 Å.

As we insert the metal ions into the protein without any structural optimization, we have to take care of strong overlap:

- we check that the overlap of the metal ion with all the atoms of the protein is less than a predefined threshold.
- we check that all possible active sites have a minimal contribution to the solvent accessible surface, where the probe sphere radius equals the metal ion radius. This surface approximately corresponds to the positions where the metal ion is placed in case of an active binding site.

With an overlap threshold of 50% and an SAS threshold of 45 Å^2, we find 70 out of 98 potential active sites for BSA.

8.3.3.4 Protein data set

The simulation protocol which has been presented in the last sections has been applied to a set of four negatively charged proteins in a solution of increasing YCl_3 concentration by M. Ziller.

[3] Please note that, similar to protonation computations, we always attach the Y^{3+}, but de-charge it in the neutral state, see Eq. (8.18).

Tab. 8.3: Protein data set and input quantities for the electrostatic calculations (H++/APBS).

Protein	PDB-entry	$Q_{\text{ini}}^{exp}\,/\,e$	$Q_{\text{ini}}^{H++}\,/\,e$	$\varepsilon_{in}/\varepsilon_{out}$
BSA	—	-10	-11	2/78.5
HSA	1N5U	-12	- 9	2/78.5
OVA	1OVA	-12	- 8	6/78.5
BLG	1BEB	-10	-10	6/78.5

Tab. 8.3 gives an overview on the PDB files used. Further, it gives information on the total charge, Q_{ini}^{exp} of the proteins at neutral pH, which has been used to guide for the choice of the inner dielectric response. The total charge estimations are taken from the following literature [25,38,42,96,110,127]. The total charge that is predicted by H++ for the given dielectric responses ε_{in} and ε_{out} is shown in the last column of Tab. 8.3. Here, we have to note that, even with a variation of the inner dielectric response $\varepsilon_{in}/\varepsilon_0$ in a range of 2 - 6, the best agreement for ovalbumin (OVA) differs from the experimental observation by $3\,e$. This clearly shows that in view of further applications, more information on the experimentally realized initial protonation state of the proteins is required.

8.3.4 Results

In this section, we present the simulation results for BSA with YCl_3. In order to correctly interpret the results, we have to consider that the binding site occupancies are determined as a *function* of the Yttrium concentration. However, the input parameter for the simulations, $\boldsymbol{pK}_m^{intr}$ and W_{ij}, also depend on the Yttrium concentration due to the ionic screening considered in the linearized Poisson-Boltzmann equation. These parameters, though, can only be calculated for a constant salt concentration. This means that the results of one particular simulation are valid only in a small region, where the Yttrium concentration used for the computation of the binding site occupancies corresponds to the given ionic strength. The influence of the ionic screening on the simulation results is discussed in Section 8.3.4.1.

In Fig. 8.8, the binding curves of all active sites of BSA are illustrated: a successive adding of salt forces the counter-ions to adsorb to the previously defined specific sites. For the binding curves shown in Fig. 8.8, the electrostatic calculations have been performed with a constant ionic strength based on $YCl_3 = 10^{-4}$ mM. This implies that the curves can be interpreted in this concentration regime, only.

8.3.4.1 Effective charge and influence of the ionic screening

Besides their role of specific binding, the metal ions together with the negative co-ions also screen the effective charge cloud of the protein-metal complex. To account for both effects in pM-dependent simulations, we calculate the charge curves for several *constant* ion concentrations (thin lines in Fig. 8.9).

Each charge curve is based on a constant ion concentration and therefore is valid in a small region around this concentration (see legend and points in Fig. 8.9). A global binding curve for the metal ion-protein binding is retrieved by interpolation of the stepwise correct charge curves. This so called *effective total charge* of the protein as a function of salt concentration is shown in Fig. 8.9 by the thick line.

We are primarily interested in the region where the effective charge vanishes, as in this regime, the electrostatic monopole forces between individual proteins get small and attractive forces determine

the behavior allowing the aggregation. We define the critical metal ion concentration, $c_{Me,crit}$, as the concentration, which is needed to neutralize a single protein-ion complex. $c_{Me,crit}$ is determined from the effective charge curves:

$$c_{Me,crit}: \quad Q_{\text{eff}}(c_{Me,crit}) = 0 \tag{8.24}$$

For BSA, the critical metal ion concentration $c_{Me,crit}$ is 0.003 mM. This corresponds to the binding of ~ 4 metal ions, since the initial total charge is $-11\,e$, see Tab. 8.3. It follows from Fig. 8.9 that in this concentration regime the screening does *not* affect the result.

8.3.4.2 Saturation effect

In protonation simulations, a saturation of the charge state usually takes place above a certain proton concentration. In our simulation, such a saturation occurs at high salt concentration, resulting in unreasonably high effective total charges. This can be already seen in Fig. 8.9, where the effective charge of the protein-metal ion complex reaches values of $+40\,e$ for pM=-2. The reason is the choice of the binding sites and our neglect of steric effects: we decided to take "almost all" surface exposed Asp and Glu side chains as possible binding sites to offer the largest possible number of degrees of freedom for the metal ions to bind, see Section 8.3.3.3. For high salt concentration, Eq. (8.20) will finally favor the bound state of all sites, even though there are high electrostatic repulsion effects. Since we are interested in the early stage of the charging process, i.e., in the concentration region where the protein exceeds an effective zero charge, this does not pose a problem in our interpretation.

8.3.5 Discussion

From the numerical simulations, we obtain information about the charge distribution of a *single* protein, whereas in the experiments we always measure the effect of a protein solution. To overcome this problem, we now propose a simple approach to transfer the previously discussed simulation results to a protein solution.

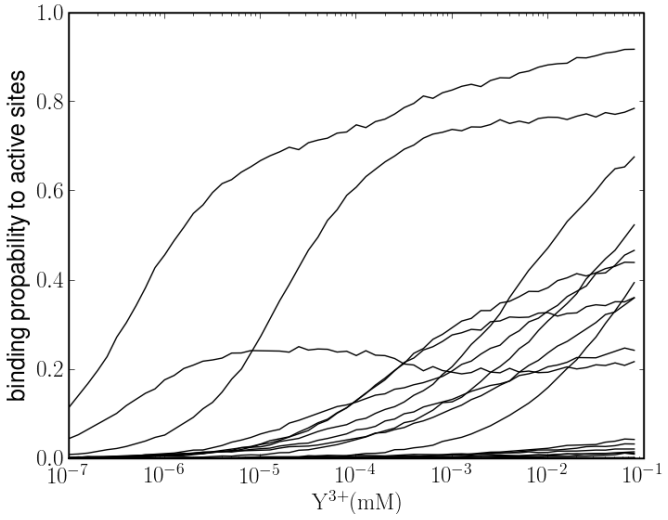


Fig. 8.8: Binding curves of BSA with YCl_3 calculated for constant ionic strength given by $YCl_3 = 10^{-4}$ mM.

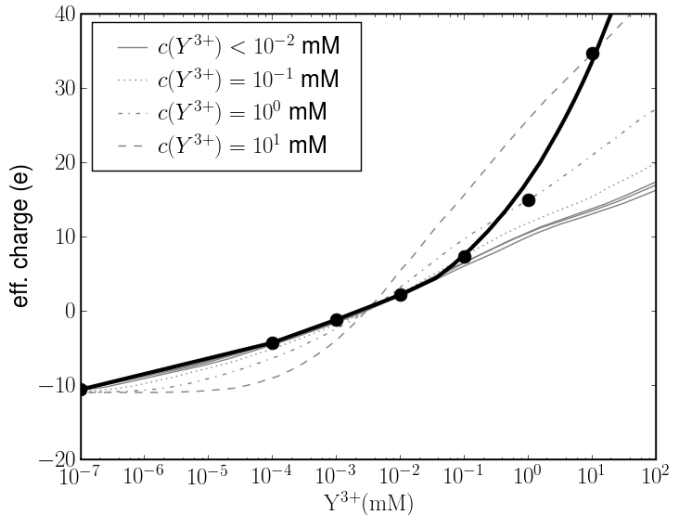


Fig. 8.9: Effective charge of BSA for increasing YCl_3 concentration.

By means of this simple theory, we can start a qualitative comparison of the experimental and the numerical results: we discuss the principal phase behavior and the course of the effective charge of the proteins to shed light on whether the theoretical approach is reasonable or not.

8.3.5.1 From a single protein to protein solution

To quantify the effect of counter-ion binding in protein solution and to give a first validation of the theoretical model, we need an estimation of the *critical protein concentration*, $c_{P,crit}$, for which the simulation results are still valid. To this end, we assume that the proteins are uncorrelated and therefore behave as independent entities as long as the protein-protein interaction energy is less than or equal to their thermal energy. In the beginning of the simulations, the proteins have a strong, negative total charge and therefore, we suggest the interaction energy to be dominated by the electrostatic part:

$$k_B T = W_{elec}^{Q_{\text{ini}}}(P_i, P_j) \tag{8.25}$$

$W_{elec}^{Q_{\text{ini}}}(P_i, P_j)$ defines the electrostatic interaction energy of two proteins i and j both with negative, total charge Q_{ini}. Accurately estimating the detailed electrostatic interaction energy in the presence of solvent, ions, and other copies of the protein is a very complex task that depends on a variety of parameters, some of which are not easily available. We thus approximate Eq. (8.25) by

$$\alpha\, k_B T \doteq \frac{Q_{\text{ini}}^2}{4\pi\varepsilon_{out} d_{eq}}\,. \tag{8.26}$$

Q_{ini} denotes the initial charge of the proteins, ε_{out} the dielectric response of water, and d_{eq} defines the equilibrium distance of two arbitrary proteins in the solution. Eq. (8.26) can be interpreted as the first term in a Taylor series of $W_{elec}^{Q_{\text{ini}}}(P_i, P_j)$ and the parameter α is used to compensate the simplicity of the interaction model, where finite size effects and the complex shape as well as the inhomogeneous charge distribution of the proteins are not taken into account. With Eq. (8.26) we define

$$c_{P,crit} := \frac{1}{4/3\pi(d_{eq}/2)^3}\frac{1}{\text{Å}^3}\,, \tag{8.27}$$

as the critical protein concentration corresponding to an average protein separation of $\frac{1}{4/3\pi(d_{eq}/2)^3}\frac{1}{\text{Å}^3}$. This implies that the effective charge curves, which originally have simulated a single protein, are valid for a solution with protein concentration $c_{P,crit}$. Up to this concentration, we assume protein-protein interactions to be negligible.

8.3.5.2 Comparison of experimental and numerical results for the first phase transition

The basic idea of the theoretical model is that the first phase transition occurs close to the curve $c^*_{Me}(c_P)$ in the phase diagram, where the protein-ion agglomerates have an effective charge of zero. Then, the overall electrostatic repulsive forces are small and a condensation can be induced by attractive forces. The simplest approach for $c^*_{Me}(c_P)$ is a linear dependence of c_P. The linear function accounts for the quantitative binding of the counter-ions to the protein's surface:

$$0 = Q_{\text{eff}}(c_P, c^*_{Me} := c_P \frac{c_{Me,crit}}{c_{P,crit}}) \,, \tag{8.28}$$

where

$$m^*_{sim} := \frac{c_{Me,crit}}{c_{P,crit}} \tag{8.29}$$

is a characteristic measure for the number of ions condensed on the surface to reveal the zero effective charge state. The curve of zero effective charge in the phase diagram is then given by

$$c^*_{Me}(c_P) = m^*_{sim} c_P \,. \tag{8.30}$$

This equation qualitatively differs from Eq. (8.16) by a y-intercept of zero. However, with the experimental data given in Tab. 8.4, we see that this agrees well with c_1 for every protein in the data set.

From the experimental point of view, $c_1 = 0$ can be interpreted in the following way: in the first phase, where the proteins are still in solution, all metal ions directly bind equally to the proteins and will not act as free ions in solution.

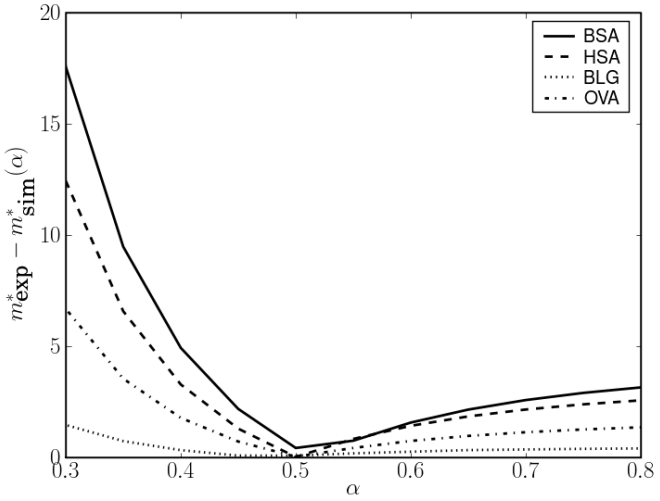


Fig. 8.10: Determination of fit parameter α.

In order to transfer the simulation results to a protein solution, we introduced a fit parameter

α, in Section 8.3.5.1. This parameter is now optimized so that m^*_{sim} fits to the experimentally observed m^*_{exp}. Fig. 8.10 shows a minimum for $\alpha = 1/2$ for the complete data set. This value is taken to define $c_{P,crit}$.

Tab. 8.4: Experimental and numerical results of the first phase transition.

	experimental results		numerical results		
Protein	m^*	c_1 [mM]	$c_{P,crit}$ [mg/ml]	$c_{Y^{3+},crit}$ [mM]	m^*_{sim}
BSA	4.30±0.5	-0.2 ±0.2	0.043	0.00290	4.83
HSA	3.40±0.5	0.1 ±0.2	0.138	0.00770	3.79
OVA	1.80±0.2	0.1 ±0.2	0.179	0.00760	1.82
BLG	0.50±0.2	0.05±0.02	0.040	0.00046	0.42

The experimentally measured as well as theoretically predicted data for m^* are listed in Tab. 8.4. With only one optimized parameter, our simulations predict m^* with good accuracy for all proteins. Even though the simplicity of the model (neglect of protein-protein interaction and simple estimation of the critical protein concentration) confines the margin of interpretation, this excellent agreement strongly supports that the determining factor of the first phase transition is a specific, quantitative binding, which results in a charge inversion of the solvated proteins. Moreover, it clarifies that the attractive force, which is finally responsible for the condensation, does not become relevant until the overall electrostatic forces are essentially neutralized.

8.3.5.3 Zeta potential comparison

By electrophoretic mobility experiments, the zeta potential of the proteins in the test set have been determined for constant protein concentration as a function of salt concentration [151]. In the following, we compare qualitatively and quantitatively the effective charge of the BSA-Y^{3+} complex with corresponding zeta potential measurements. Such a comparison is interesting as - at least in colloidal systems - the zeta potential indicates the movement of the ion-macromolecule complex. In this way, it is used as a measure of the effective surface charge in ionic systems. Of course, the charge distribution of proteins is more complex than for colloids. However, as a first analogy of the surface charge of colloids, the effective total charge of the protein is certainly a reasonable quantity.

Fig. 8.11 shows zeta potential measurements for varying YCl_3 concentration for two different BSA concentrations (2mg/ml, 5mg/ml). Assuming a near-quantitative binding of the counter-ions on the surface of every single protein, we can adjust the salt concentration of the original charge curve to correspond to the protein concentrations used in the experimental setup. This means that we assume a linear relationship between the salt and the protein concentration for every *constant* effective charge state,

$$Q_{\text{eff}}(c_P, c_P \frac{c_{Me}}{c_{P,crit}}) = Q_{\text{eff}}(c_{P,crit}, c_{Me}) \,, \tag{8.31}$$

where in particular, for the couple ($c_P = c_{P,crit}$, $c_{Me} = c_{Me,crit}$) we reach the zero effective charge state. In Fig. 8.11, we illustrate the scaled charge curves for the two concentrations of BSA, which are used in the zeta potential measurements. Both, the measured zeta potential and the simulated effective protein charge, become zero at nearly the same counter-ion concentration. The zero point depends on the protein concentration: an increase in protein concentration is accompanied by the need of a higher counter-ion concentration and this results in the shift observed in Fig. 8.11. In the presented theory, such a shift is directly implied by the scaling of the effective charge curve

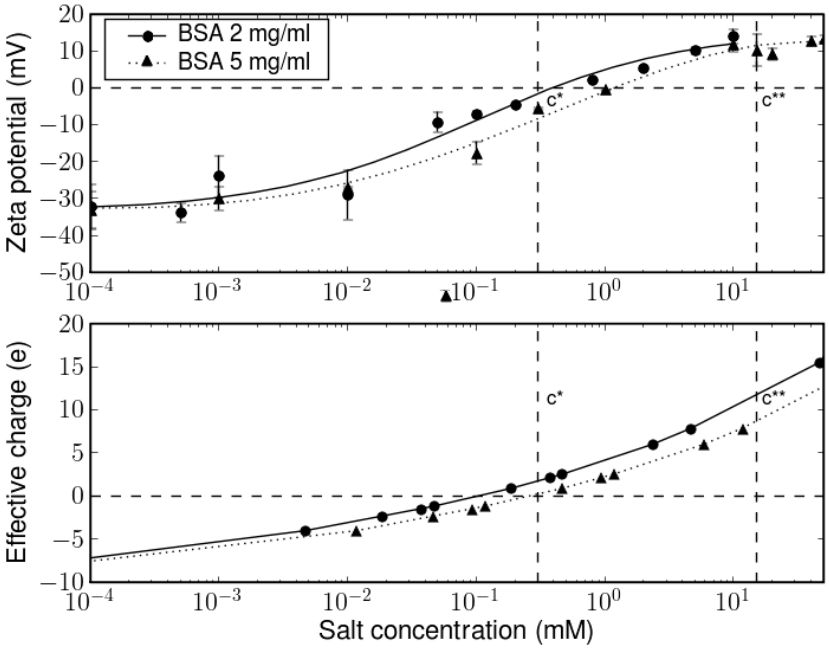


Fig. 8.11: Comparison of zeta potential (exp) and effective charge (sim) of BSA. The vertical, dashed lines indicate the border of the condensed phase for 5mg/ml BSA concentration, which ranges from $\sim$0.1 to $\sim$25 mM salt concentration.

given in Eq. (8.31) and is thus reproduced by our simulations as well. Close to the transition into the condensed phase, both, the zeta potential and the effective charge of the proteins, cross zero. In Section 8.3.1, this experimental result has already been discussed as a supporting factor for the assumption that the charge inversion determines the first phase transition.

The zeta potentials and the effective protein charge curves exhibit similar shapes. Here, we note that for the zeta potential one clearly sees a saturation effect when the counter-ion concentration is increased. As explained in Section 8.3.4.2, the saturation in the numerical studies takes place at higher salt concentrations due to deficiencies of the theoretical model for high metal ion concentrations and therefore cannot be seen.

8.3.6 Conclusion

In the last sections, we proposed and applied a modified titration model, which explicitly incorporates the binding of metal ions to localized amino acid side chains (c.f. Section 8.3) in order to account for the highly interesting effect of reentrant condensation in protein solution. With this model, we predicted the bound ion distribution on the surface of an isolated protein. The low metal ion concentration together with the high binding affinity justifies the assumption of a near-quantitative binding of the metal ions to the surface exposed acidic side chains of every protein and this implicitly explains the linear shape of the first phase transition curve. With only one optimized parameter, the numerically determined curves of charge inversion lie next to the experimentally observed first phase transition for a set of four analyzed proteins. Experimentally measured zeta potentials agree well with the effective charge curves and therefore support the theoretical concept of specific metal ion binding.

Moreover, our model qualitatively explains the second phase transition, i.e., the dissolving of

the protein from the condensed phase: the increase of the counter-ion concentration above c^* causes the counter-ions to further bind to free active sites. The aggregates finally re-dissolve, when m^{**} counter-ions are bound and the overall interaction becomes electrostatically repulsive again. A quantitative prediction of m^{**} and c_2 remains a challenge, since ion-ion correlations as well as protein-protein interactions are not included in the current theoretical model. To take those interactions into account, we have to know more about the balance of attractive and repulsive forces in the condensed phase.

To simulate the protein-metal ion binding process and to transfer the numerical results to the experimental data, we made a number of approximations and assumptions, which certainly limit the range of interpretation. Despite its simplicity, the proposed theory can consistently explain *all* the experimental observations going along with the first phase transition of reentrant condensation. Therefore, it is the first theoretical model, which can reasonably elucidate this highly complex phenomenon.

Chapter 9

Summary and outlook

This thesis focuses on the theoretical modeling and the deeper understanding of two types of electrostatic interactions in biomolecular systems: the main part of this work aims at describing the water-water correlations due to the hydrogen bond network, whereas the second part is concerned with the electrostatic effect of additionally solvated ions.

By recalling the electrostatic material equations and its adaption to biomolecular systems in the Chapters 2-3, we provided a self-sufficient background for all following considerations.

In Chapter 3, we further re-investigated and partly corrected the derivations of nonlocal electrostatics done in [52]. Additionally, we included the ideas given in [39] by the introduction of an additional, physical field, the so called correlation field, which captures the contribution to the polarization field due to the nonlocal water correlations (c.f. Section 3.3.2). Although this field has already been defined in previous work, its physical meaning, which of course is important to correctly model its influence, has not been investigated yet. In Chapter 3, we gave for the first time an interpretation of the correlation field. Within this study, we realized that the already established Lorentzian water models [33, 39, 52] lack in explicitly incorporating the physical influence of the biomolecular surface as a disturbing factor for the water network. The crucial impact of dielectric boundaries or solid-liquid interfaces on the water correlations, however, is experimentally proven and among others becomes evident as surface tension or first-shell effects [20, 51].

In Chapter 4, we proposed an *extended* integral representation of the dielectric operator. The extension consists of additional boundary value integrals, whose physical meaning is apparent in the corresponding differential formulation: they determine the boundary conditions of the correlation field on the molecular surface (c.f. Section 4.3). With this theoretical formulation, we go beyond the previous models [39, 52] and further shed light on their physical approximations: the so called Newton Vector model (NVM) is the model presented in [39]. Therein, the correlation field is directly deduced from the original Lorentzian model [33], which describes water without any disturbances due to dielectric boundaries. Thus, the NVM ignores a possible change of the correlation field on the molecular surface (c.f. Section 4.3.1). The so called Dirichlet Vector model (DVM) is the vectorial analog of the scalar model discussed in [52]. It constitutes another extreme: because of the spatial constraints and the high electrostatic potential, the correlation field is set to zero on the molecular surface (c.f. Section 4.3.2), which corresponds to a maximization of the hydrogen correlations.

Besides the deeper understanding of these already established nonlocal models, this novel formulation opens a new dimension for a successful modeling of water correlations in biomolecular applications, accounting for the biomolecular surface as a disturbing or supporting factor of the hydrogen bond formation.

The general integral formulation reveals another interesting aspect concerning the field energy of the biomolecular system: the additional boundary integrals cause an additional energy contribution. Since the boundary integrals determine the behavior of the water correlations on the molecular surface, we interpret these energy contribution as a measure for the energy comprised

in the reorientation and breaking of hydrogen bonds (c.f. Section 4.3). In continuum theories, where the water's structural effects are not further considered, such *entropic cavity* terms have to be explicitly added (c.f. Section 3.1.4.2).

In prospect of a broader application of nonlocal electrostatics, models based on a gradient formulation are, if not necessary, at least preferable, as they reduce the high demand on memory and runtime of the solution procedure for the vectorial equations [39]. Thus, based on the two vectorial models, the NVM and the DVM introduced above, we physically motivated two approximated models, the Dirichlet Scalar model (DSM) and the Newton Scalar model (NSM) in Chapter 5. Both deal with a gradient ansatz of the vectorial nonlocal fields. The similarity of the potentials of the vectorial and the scalar nonlocal models in the spherically symmetric case encouraged us to consider these scalar models in further studies (c.f. Section 5.3).

Having provided a novel nonlocal formulation and having derived a set of different, nonlocal models, Chapter 6 is dedicated to developing a numerical tool to easily implement and solve the proposed as well as future models of nonlocal electrostatics. The method of choice was the Explicit Jump Immersed Interface method (EJIIM). On the one hand, this method offers all advantages of finite difference schemes such as an easy extensibility and the possibility to use efficient numerical schemes to solve the algebraic equations. On the other hand, it yields highly accurate results (c.f. Section 6.2). Indeed, the EJIIM lives up to its promise as could be demonstrated by a detailed, numerical comparison with the DSM potentials generated by the Boundary Element method (BEM) (c.f. Section 6.4).

The exterior boundary condition, which is a necessary input for the EJIIM, is provided by the potentials of the charge distribution directly immersed in the solvent (c.f. Section 6.2.4). The theoretical considerations as well as the numerical results impressively demonstrated that these boundary conditions do not spoil the accuracy of the solution. Indeed, for the physically reasonable choice of a constant inner dielectric response $\varepsilon_\Omega = 2\varepsilon_0$, the study of the analytical boundary approximations for the DSM reveals that its evaluation in the whole outer region yields an excellent estimation of the numerical solution (c.f. Section 6.4.2 and Section 7.2.1.1). This means that we found a fast analytical approximation of the electrostatic potential ϕ_Σ in the DSM.

The accuracy of the EJIIM relies further on an adequate description of the molecular surface in a 3-dimensional Cartesian grid. We developed and implemented a numerical tool to generate the grid-based SES, SAS, and VdW surface information of an arbitrary molecule. The algorithm is based on the fact that all of these surface models can be constructed by an union of spheres (c.f. Section 6.3). Moreover, we combined this algorithm with a marching cubes algorithm in order to generate surface triangulations, which - after an automated coarsening step - are suitable as input for the BEM (c.f. Section 6.3.5). Thus, our algorithm for the grid-based surface generation opens not only the possibility for a fully automated finite difference solver but eventually paves the way for an automated BEM to compute nonlocal electrostatic potentials of biomolecules.

Besides extensive studies of the BEM and the EJIIM in terms of accuracy and convergence, we solved the nonlocal electrostatic equations for various proteins in Chapter 7. A comparison of the different nonlocal and local electrostatic potentials reveals that the shielding effect of the water network causing a higher outer electrostatic potential is predicted in *all* nonlocal models. Moreover, all these models predict similar electrostatic field energies. We conclude that in order to capture the nonlocal features in electrostatic theories, i.e., in order to go beyond the local electrostatics, all the proposed models are appropriate as starting point.

A closer look on the differences of the nonlocal models, however, reveals that on average the NVM yields a higher potential near the molecular surface than the other nonlocal models. While for small molecules with only a few buried charges, the variations in magnitude are marginal (c.f. Section 7.1), they are most obvious for molecules with a charge distribution next to the dielectric

boundary. In particular, at the entrance area of the binding pocket of trypsin, the electrostatics sensitively hinges on the chosen nonlocal model (c.f. Section 7.2). We interpreted this finding in terms of the different behavior of the correlation field on the molecular surface (c.f. Section 7.2.1.3). vanishing correlation field in the DSM.

Concerning these water correlation studies we now draw a conclusion, which comprises theoretical ideas and application projects for future progress.

First, we conclude that in comparison to local electrostatics all the nonlocal models reveal the same order of magnitude in the electrostatic potential. This means that the overall nonlocal features are well represented and stable against slight variations of the boundary condition. In this respect, it is reasonable to pursue the nonlocal scalar approaches which are feasible for a numerical solving procedure.

If the aim of a study is the detailed and accurate modeling of the potential on the molecular surface, we found that boundary effects become important. This holds, in particular, when the charge distribution becomes complex as well as when the molecule's extension exceeds the correlation length of water. A decision on the question which theoretical model is most suitable in representing realistic potentials, however, cannot be given on the basis of the presented comparison. All the implemented models do not yet fully exploit the new formulation as they enforce a *globally uniform* nonlocal behavior on the molecular surface: the Newton models enforce an undisturbed, whereas the Dirichlet models enforce a vanishing correlation field. With the new formulation, we gain the possibility to account for variations of the correlation field on the molecular surface in a nonetheless *continuum* description. Local variations of the dielectric response can, for instance, be caused by regions of high polarity or by surface exposed amino acids which are able to bind water molecules due to hydrogen bonds. Thus, future work lies in the careful study of experimental data that reveal the structural features of water within the first and second solvation shell. The hydrogen bonds are, for example, studied in [51], where THz spectroscopy provides new insight "in the collective dynamics of water molecules on a spatial scale spanning several solvation shells". Besides experimental data, ab initio molecular dynamic simulations allow for a detailed time and space resolution of the molecular system, which makes them attractive guides for a conception of boundary conditions in our framework.

Second, with the new theoretical and numerical framework, we are now in the position to develop a fully automatic, robust solver for various, (non)local electrostatic models freely available in a biomolecular framework. This is necessary to make the nonlocal model available for the scientific community. In order to achieve this aim, the current Matlab implementation of the EJIIM has to be transferred to `C++` and combined with its - up to now separate - input generator. Simultaneous to this task, a range of technical details can easily be incorporated to speed up the method, for instance, a parallelization of the merging procedures for the grid-based surface generation as well as the implementation of multigrid techniques for an iterative, fast and efficient solving procedure. Further, with the accurate grid-based surface information, we offer the possibility for an automatic BEM for (non)local electrostatics of biomolecules, as well. The only requirement is the integration of a mesh coarsening procedure, which can be applied on the triangulation originally generated by the grid-based surface generator. A suitable algorithm could be the Quadric Edge Collapse Decimation algorithm by M. Garland and P. Heckbert [44].

Besides the aim to directly solve the nonlocal electrostatic model equations, it is often desirable to have a fast and efficient estimation of the electrostatics within a molecular system. As discussed above, we proposed an analytical approximation of the electrostatic potential ϕ_Σ in the DSM for the parameter regime $\varepsilon_\Omega \approx \varepsilon_\infty$. This means that instead of solving the complex system of differential equations, we found an extremely efficient approximation of the outer potential predicted by the DSM. It captures the physics of nonlocal electrostatics, while improving the speed of the

calculations. In this way, it is comparable with the role of the Generalized Born model as fast approximation of the electrostatic potential within the Poisson-Boltzmann theory. Incorporating this fast prediction within a biomolecular software package offers attractive applications. For instance, for the first time it enables to study the dynamics of ligands searching for the protein's binding pocket within the nonlocal theory of electrostatics.

A task which is essential for further application of nonlocal electrostatics to biomolecular research, such as solvation studies of biomolecules in their natural surrounding, is the design of a parameter set of the atom type radii optimized for the nonlocal response. Commonly used force fields [48, 84, 98, 116, 125] are based on parameter sets, which work well using a local electrostatic contribution. With a "nonlocal parameter set" at hand, physically meaningful quantities, such as the solvation energy of molecules, but also the binding affinities of small ligands to the protein's binding pocket, can be studied and compared to available experimental data. We mentioned above that the general nonlocal theory provides contributions to the electrostatic field energy which originate from the mechanical work needed for the cavity formation, i.e., water molecules are urged to make room for the biomolecule, implying a reorientation of the hydrogen bonds. The study of these energy terms has to accompany the parameter optimization as it promises a fruitful unification of entropic and electrostatic energy terms.

In Chapter 8 we extend our discussion to ionic solutions.

First, we succeeded in describing the proteins in ionic solutions of correlated water molecules, i.e., we combined the linearized Poisson-Boltzmann theory with a nonlocal water model (c.f. Section 8.2). In the differential formalism, the ionic screening is mediated by a Yukawa operator. This is an interesting result, as the nonlocal shielding of the correlated water molecules is characterized by a Yukawa operator as well. The detailed discussion of the nonlocal correlation and the ionic screening effects on the spherically symmetric case revealed that the ion screening dominates the overall electrostatics and suppresses the shielding effect of correlated water. In order to demonstrate the possibilities of the EJIIM we extended the DSM to account for the screening effect of small ions. First applications of the extended nonlocal model on small molecules support that the ionic screening strongly weakens the electrostatic potential in the regime of realistic Debye screening and water correlation lengths. Questions such as finite size effects of the ions as well as the dependence of the water correlations on the salt concentration have only been touched, so that further studies must follow in order to clarify the impact of these approximations. However, this study clearly demonstrates that, by means of the newly developed finite difference method, we can address such interesting and relevant problems in the future.

Furthermore, in Chapter 8, we focused on another type of correlations, namely the coordinative binding of metal ions to the protein surface. This specific interaction is responsible for a range of interesting and important effects in protein solutions, such as the *reentrant condensation* of proteins (c.f. Section 8.3.1). In order to qualitatively explain this recently discovered effect, we proposed a simple, heuristic model that considers the specific metal ion binding by attributing a binding affinity constant for metal ions to negatively charged, surface exposed amino acid side chains (c.f. Section 8.3.2). This idea reminds of the theoretical treatment of titration experiments, where the binding and the release of hydrogens are considered by dissociation constants which depend on the type of amino acid as well as on the protein environment [111]. Because of this analogy, we implemented a modified titration program that predicts the ion distribution on the surface of an isolated protein as a function of metal ion concentration.

With the assumption of a near-quantitative binding of the metal ions to the protein's surface, the linear shape of the first transition curve is implicitly explained. As the simulation predicts the distribution of metal ions on the surface of a single protein, we proposed a simple estimation of the protein concentration for which the simulation results are valid. With only one optimized

parameter, the numerically determined curves of charge inversion lie next to the experimentally observed first phase transitions for a set of four proteins. A further support for the proposed theory of specific binding is given by a qualitative agreement of the shape of the binding curves and the experimentally measured zeta potentials.

With the beginning of the condensed phase, attractive protein-protein interactions cause the protein to aggregate. In order to incorporate such interactions and to finally make predictions for the condensed phase as well as for the second phase transition curve, the simulations have to be extended to protein solutions. An appropriate Monte Carlo model, which is capable to account for both, the specific correlations between protein and metal ions and protein-protein interactions, is described by Kesvatera and coworkers [68] with respect to protonation theory. The effects of multivalent ions in biomolecular systems are highly complex and can also originate from ion-ion correlations [46]. Although our simple theory qualitatively explains the characteristics of reentrant condensation in protein solutions, a focus of future work must also be the inclusion of such ion-ion correlations in the theoretical framework in order to assess their effect.

In summary, we pursued different directions, namely theoretical studies, the development of new numerical tools as well as model extensions, to progress in the understanding of correlation induced electrostatic phenomena and their treatment in biomolecular systems.

In our opinion, the theoretical results of this thesis have shown that the theory of nonlocal electrostatics unifies important aspects relevant in biomolecular systems. This is, on the one hand, the overall electrostatics of the solute and the solvent and, on the other hand, first shell effects of the water network around the molecule. In this way, the transition from the currently established "homogeneous continuum" approximation to what can be called a "structured continuum" - a continuum that is not blind to correlations among its "constituents" - might imply a progress which might drastically change our current understanding of many electrostatically dominated processes on a molecular scale.

The numerical tools and the analytical approximations we developed for the water correlations and for the specific ion-binding complete the theoretical ideas of this thesis. They are innovative tools and necessary prerequisites for further studies of correlation phenomena in biomolecular systems.

We hope that the work presented in this thesis - the incorporation of specific boundary effects into the equations of nonlocal electrostatics, the development of an efficient and accurate finite difference solver, and the numerical description of an ion-binding process - serves as a contribution to a deeper understanding of correlation induced electrostatical effects.

Chapter 10

Appendix

10.1 Notations relating to spaces of integrable functions

For a modern theory of partial differential equations, the classical concepts of continuity and differentiability turn out to be insufficient, since they usually demand too much regularity of the functions we want to study. This can, for instance, be seen on the transmission conditions in electrostatic problems of different dielectrics derived in Section 2.2.2: we found that the normal component of the electrostatic field $\boldsymbol{E}$ has a *discontinuous jump* on the interface Γ of the two different dielectrics and thus, the electrostatic field is not differentiable on Γ. The Sobolev spaces are the modern replacement for classical spaces, such as the space of continuously differentiable functions $\mathcal{C}^1$, in which to look for solutions of partial differential equations.

In the following, we want to give a brief definition of the Sobolev spaces we introduced in Section 4.2 to define solutions of boundary and transmission problems of the Yukawa operator. As the mathematical theory is beyond the scope of this work, we refer to [27, 52, 90, 121] for an extended introduction and a complete definition of *weak differentiability* and Sobolev spaces.

10.1.1 Sobolev spaces for open subsets Ω of $\mathbb{R}^d$

Definition 10.1 A function f on Ω is stated to be *locally integrable* in Ω, if f is integrable over any compact subset $\tau \subset \Omega$. The space of integrable functions on Ω is denoted by $L^1_{loc}(\Omega)$.

Definition 10.2 A locally integrable function $f \in L^1_{loc}(\Omega)$ is said to possess a generalized or *weak derivative* with respect to $\boldsymbol{x}$ if there exists a locally integrable function $v \in L^1_{loc}(\Omega)$, with

$$\int_\Omega v(\boldsymbol{x})\phi(\boldsymbol{x})d\boldsymbol{x} = -\int_\Omega f(\boldsymbol{x})\partial_{\boldsymbol{x}}\phi(\boldsymbol{x})d\boldsymbol{x}\,,$$

for all infinitely-differentiable function ϕ with compact support in Ω, i.e., $\phi \in \overline{C^\infty}(\Omega)$. In that case, we say that

$$\partial_{\boldsymbol{x}} f(\boldsymbol{x}) := v(\boldsymbol{x})$$

is the generalized derivative of $f(\boldsymbol{x})$. Iterating this definition, we can define partial derivatives ∂^α of arbitrary positive order α. This definition is motivated by the integration technique of "Integration by parts".

Definition 10.3 Let Ω be an open subset of $\mathbb{R}^d$. For $k \in \mathbb{N}_0$, the Sobolev space $W^k_2(\Omega)$ is defined by

$$W^k_2(\Omega) := \{u \in L^2(\Omega) : \partial^\alpha u \in L^2(\Omega) \text{ for } |\alpha| \leq k\}\,,$$

where $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}^d_0$. $|\alpha| = \alpha_1 + \cdots + \alpha_d$, and $\partial^\alpha u(\boldsymbol{x}) = \partial^\alpha_1 \ldots \partial^\alpha_d u(\boldsymbol{x})$ are to be understood as weak partial derivatives.

The Sobolev space $W_2^k(\Omega)$ is equipped with the norm

$$||u||_{W_2^k(\Omega)} := \left(\sum_{|\alpha| \leq k} \int_\Omega |\partial^\alpha u(\boldsymbol{x})|^2 dx \right)^{\frac{1}{2}}$$

and it is a Hilbert space equipped with inner product

$$(u, v)_{W_2^k(\Omega)} := \sum_{|\alpha| \leq k} \int_\Omega \partial^\alpha u(\boldsymbol{x}) (\partial^\alpha v(\boldsymbol{x}))^* d\boldsymbol{x} \,.$$

The definition of Sobolev spaces $W_2^l(\Omega)$ can be extended for any arbitrary $l > 0$:

Definition 10.4 Let Ω be an open subset of $\mathbb{R}^d$. For $s = k + \mu$ with $k \in \mathbb{N}_0$ and $\mu \in (0, 1)$, the Sobolev space $W_2^s(\Omega)$ is defined as

$$W_2^s(\Omega) = \{u \in W_2^k(\Omega) : |\partial^\alpha u|_{\mu,\Omega} < \infty \text{ for } |\alpha| = k\} \,,$$

where the Sobolev-Slobodeckii semi-norm $|\cdot|_{\mu,\Omega}$ is given as

$$|u|_{\mu,\Omega} := \left(\int_\Omega \int_\Omega \frac{|u(\boldsymbol{x}) - u(\boldsymbol{y})|^2}{|\boldsymbol{x} - \boldsymbol{y}|^{d+2\mu}} \, d\boldsymbol{x} \, d\boldsymbol{y} \right)^{\frac{1}{2}} .$$

Again, $W_2^s(\Omega)$ is a Hilbert space with respect to the inner product

$$(u, v)_{W_2^s(\Omega)} := (u, v)_{W_2^k(\Omega)} + \sum_{|\alpha|=k} \int_\Omega \int_\Omega \frac{[\partial^\alpha u(\boldsymbol{x}) - \partial^\alpha u(\boldsymbol{y})][\partial^\alpha v(\boldsymbol{x}) - \partial^\alpha v(\boldsymbol{y})]}{|\boldsymbol{x} - \boldsymbol{y}|^{d+2\mu}} \, d\boldsymbol{x} \, d\boldsymbol{y}$$

A second family of Sobolev spaces $H^s(\mathbb{R}^d)$ can be introduced by using the Fourier transform

$$\tilde{u}(\xi) = \int_{\mathbb{R}^d} e^{-i2\pi \, \boldsymbol{x} \cdot \xi} u(\boldsymbol{x}) d\boldsymbol{x}$$

for $u \in L^1(\Omega)$. The Sobolev space $H^s(\mathbb{R}^d)$ for $s \in \mathbb{R}$ is defined by

$$H^s(\mathbb{R}^d) := \{u \in \mathcal{S}'(\mathbb{R}^d) : \mathcal{J}^s u \in L^2(\mathbb{R}^d)\} \,,$$

where $\mathcal{S}'$ is the space of the continuous linear functionals on the Schwartz space $\mathcal{S}(\mathbb{R}^d)$ of rapidly decreasing functions in $C^\infty(\mathbb{R}^d)$,

$$\mathcal{S}(\mathbb{R}^d) := \{\phi \in C^\infty(\mathbb{R}^d) : \sup_{\boldsymbol{x} \in \mathbb{R}^d} |\boldsymbol{x}^\alpha \partial^\beta \phi(\boldsymbol{x})| < \infty \text{ for all multi-indices } \alpha \text{ and } \beta\} \,,$$

and where $\mathcal{J}^s$ is the Bessel potential of order s,

$$\mathcal{J}^s u(\boldsymbol{x}) = \int_{\mathbb{R}^d} (1 + |\xi|^2)^{\frac{s}{2}} e^{i2\pi \, \boldsymbol{x} \cdot \xi} d\xi \quad \text{for } \boldsymbol{x} \in \mathbb{R}^d \,.$$

The Sobolev spaces $H^s(\mathbb{R}^d)$ and $W_2^s(\mathbb{R}^d)$ coincide for each $s \geq 0$.

For general domains $\Omega \subset \mathbb{R}^d$, the following Sobolev spaces $H^s(\Omega)$ are defined:

$$H^s(\Omega) := \left\{ u = \tilde{u}\big|_\Omega : \tilde{u} \in H^s(\mathbb{R}^d) \right\},$$

with the norm

$$||u||_{H^s(\Omega)} := \inf_{\tilde{u} \in H^s(\mathbb{R}^d), \tilde{u}\big|_\Omega = u} ||\tilde{u}||_{H^s(\mathbb{R}^d)}.$$

Further, we define $\tilde{H}^s(\Omega)$ as the closure of $C_0^\infty(\Omega)$ under the full $H^s(\mathbb{R}^d)$-norm and $\tilde{H}_0^s(\Omega)$ as the closure of $C_0^\infty(\Omega)$ under the $H^s(\Omega)$-norm, i.e.,

$$\begin{aligned} \tilde{H}^s(\Omega) &:= \overline{C_0^\infty(\Omega)}^{||\cdot||_{H^s(\mathbb{R}^d)}}, \\ H_0^s(\Omega) &:= \overline{C_0^\infty(\Omega)}^{||\cdot||_{H^s(\Omega)}}. \end{aligned}$$

This finally allows us to give the definitions of the Sobolev spaces for $0 < s \in \mathbb{R}$:

Definition 10.5 Let $s \in \mathbb{R}$. For $0 < s$, the Sobolev spaces $H^s(\Omega)$ and $\tilde{H}^s(\Omega)$ are defined as the closures of $C^1(\Omega)$ and $C_0^1(\Omega)$ under the corresponding Sobolev-Slobodeckii norm. For $s < 0$, the Sobolev space $H^s(\Omega)$ is defined as the dual space of $H^{-s}(\Omega)$ with associated norm

$$||u||_{H^s(\Omega)} := \sum_{0 \neq v \in \tilde{H}^{-s}(\Omega)} \frac{\int_\Omega u(\boldsymbol{x}) v(\boldsymbol{x}) d\boldsymbol{x}}{||w||_{\tilde{H}^{-s}(\Omega)}}$$

Similarly, $\tilde{H}^s(\Omega)$ is defined as the dual space of $H^{-s}(\Omega)$ for $s < 0$.

10.1.2 Sobolev spaces on boundaries Γ

We assume that $\Omega \subset \mathbb{R}^d$ is a Lipschitz domain. The L^2-norm on the boundary $\Gamma = \partial\Omega$ is defined as

$$||u||_{L^2(\Gamma)} := \left(\int_\Gamma |u(\boldsymbol{x})|^2 d\boldsymbol{x}_s \right)^{\frac{1}{2}}.$$

For $s \in (1, 0)$, the Sobolev-Slobodeckii-norm is defined by

$$||u||_{H^s(\Gamma)} := \left(||u||_{L^2(\Gamma)}^2 + \int_\Gamma \int_\Gamma \frac{|u(\boldsymbol{x}) - u(\boldsymbol{y})|^2}{|\boldsymbol{x} - \boldsymbol{y}|^{d-1+2\mu}} d\boldsymbol{x}_s \, d\boldsymbol{y}_s \right)^{\frac{1}{2}}$$

Definition 10.6 Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain with boundary $\Gamma = \partial\Omega$. The spaces $L^2(\Gamma)$ and $H^s(\Gamma)$ are defined as closures,

$$\begin{aligned} L^2(\Gamma) &:= \overline{C^0(\Gamma)}^{||\cdot||_{L^2(\Gamma)}}, \\ H^s(\Gamma) &:= \overline{C^0(\Gamma)}^{||\cdot||_{H^s(\Gamma)}} \quad \text{for } s \in (0, 1). \end{aligned}$$

The spaces $L^2(\Gamma)$ and $H^s(\Gamma)$ for $s \in (0,1)$ are Hilbert spaces equipped with the inner products

$$\begin{aligned}(u,v)_{L^2(\Gamma)} &:= \int_\Gamma u(\boldsymbol{x})(v(\boldsymbol{x}))^* d\boldsymbol{x}_s\,,\\ (u,v)_{H^s(\Gamma)} &:= (u,v)_{L^2(\Gamma)} + \int_\Gamma\int_\Gamma \frac{[u(\boldsymbol{x})-u(\boldsymbol{y})][v(\boldsymbol{x})-v(\boldsymbol{y})]^*}{|\boldsymbol{x}-\boldsymbol{y}|^{d-1+2\mu}} d\boldsymbol{x}_s\, d\boldsymbol{y}_s \text{ for } s\in(0,1)\end{aligned}$$

For negative indices s, the Sobolev spaces $H^s(\Gamma)$ are defined by duality with respect to the $L^2(\Gamma)$-inner product.

10.1.3 Spaces of the Helmholtz decomposition

In Section 5.1 we introduced the so called Helmholtz decomposition, which describes the orthogonal decomposition of the vectorial space of integrable functions defined in a domain Ω, $L^2(\Omega)^3$. To this end, we introduced various function spaces which we define in the following. For an detailed discussion, the interested reader is referred to [29], in particular, Chapter IX. on p. 200.

$L^1(\Omega)$ $L^1(\Omega)$ is the space of measurable functions on Ω such that $(|f(x)|,\, x\in\Omega)$ is integrable on the open, bounded set $\Omega\in\mathbb{R}^3$, i.e.,

$$\int_\Omega |f(x)|\,dx < \infty\,.$$

This is a Banach space equipped with the norm

$$f \to \int_\Omega |f(x)|\,dx\,.$$

$L^2(\Omega)$ $L^2(\Omega)$ is the space of measurable functions on Ω such that $(|f(x)|^2,\, x\in\Omega)$ is integrable on the open, bounded set $\Omega\in\mathbb{R}^3$, i.e.,

$$\left[\int_\Omega |f(x)|^2\,dx\right]^{\frac{1}{2}} < \infty\,.$$

This is a Banach space equipped with the norm

$$f \to \left[\int_\Omega |f(x)|^2\,dx\right]^{\frac{1}{2}}\,.$$

Based on the previous definitions of $L^2(\Omega)$ and the Sobolev space $H^1(\Omega)$, we can define the three spaces in which the vector space of integrable functions can be orthogonally decomposed. In this case, Ω has to be a regular and bounded set in $\mathbb{R}^3$:

$H_0(\operatorname{curl}0,\Omega)$ $H_0(\operatorname{curl}0,\Omega) := \{\boldsymbol{u}\in L^2(\Omega)^3, \boldsymbol{\nabla}\times\boldsymbol{u}=0, \boldsymbol{u}\times\boldsymbol{n}\big|_{\partial\Omega}=0\}$

$H_0(\operatorname{div}0,\Omega)$ $H_0(\operatorname{div}0,\Omega) := \{\boldsymbol{w}\in L^2(\Omega)^3, \boldsymbol{\nabla}\cdot\boldsymbol{w}=0,\ \boldsymbol{w}\cdot\boldsymbol{n}\big|_{\partial\Omega}=0\}$

$\mathrm{L}^2_{rg}(\Omega)$ $\mathrm{L}^2_{rg}(\Omega) := \big(\operatorname{curl} H^1(\Omega)^3\big)\cap\big(\operatorname{grad} H^1(\Omega)\big)$

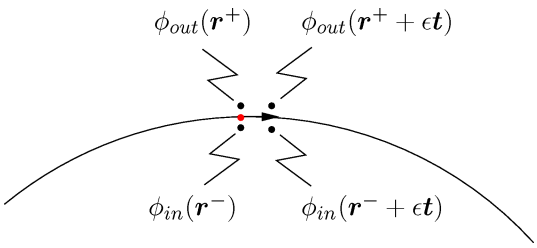


Fig. 10.1: Continuity of the tangential derivatives of the electrostatic potential on an arbitrary surface Γ implies the continuity of the potential itself on Γ. The red point on the surface represents the position $\boldsymbol{r}$ for which the derivative is calculated.

10.2 Continuity condition of the electrostatic field and its potential

The Maxwell equations tell us that the electrostatic field can be expressed - up to a gauge constant - by the gradient of a scalar potential which we denoted the electrostatic potential ϕ. In the following, we prove that the continuity of the tangential components of the electrostatic field $\boldsymbol{E}$ on an arbitrary surface $\Gamma \subset \mathbb{R}^3$ corresponds to the continuity of the electrostatic potential on this surface:

$$\begin{aligned}
& \boldsymbol{n} \times (\boldsymbol{E}_{out}(\boldsymbol{r}) - \boldsymbol{E}_{in}(\boldsymbol{r})) = \boldsymbol{0}, && \boldsymbol{r} \text{ on } \Gamma \\
\Leftrightarrow \quad & \boldsymbol{t} \cdot (\boldsymbol{E}_{out}(\boldsymbol{r}) - \boldsymbol{E}_{in}(\boldsymbol{r})) = 0 \quad \text{and} \quad \boldsymbol{l} \cdot (\boldsymbol{E}_{out}(\boldsymbol{r}) - \boldsymbol{E}_{in}(\boldsymbol{r})) = 0, && \boldsymbol{r} \text{ on } \Gamma \\
\Rightarrow \quad & \phi_{out}(\boldsymbol{r}) - \phi_{in}(\boldsymbol{r}) = 0, && \boldsymbol{r} \text{ on } \Gamma\,,
\end{aligned}$$

where $\boldsymbol{n}$ is the normal and $\{\boldsymbol{t}, \boldsymbol{l}\}$ are the tangential vectors on the surface Γ at position $\boldsymbol{r} \in \mathbb{R}^3$.

We know that the continuity of the tangential components $\boldsymbol{t}$ and $\boldsymbol{l}$ of the electrostatic field $\boldsymbol{E}$ implies the continuity of the tangential derivatives of the electrostatic potential. Considering an arbitrary tangential direction $\boldsymbol{t}$ at position $\boldsymbol{r}$ (red point in Fig. 10.1), we thus write

$$\begin{aligned}
& \boldsymbol{t} \cdot (\boldsymbol{E}_{out}(\boldsymbol{r}) - \boldsymbol{E}_{in}(\boldsymbol{r})) = 0, && \boldsymbol{r} \text{ on } \Gamma && (10.1) \\
\Leftrightarrow \quad & \boldsymbol{t} \cdot (\boldsymbol{\nabla}\phi_{out}(\boldsymbol{r}) - \boldsymbol{\nabla}\phi_{in}(\boldsymbol{r}) = 0, && \boldsymbol{r} \text{ on } \Gamma && (10.2) \\
\Leftrightarrow \quad & \partial_t\phi_{out}(\boldsymbol{r}) - \partial_t\phi_{in}(\boldsymbol{r}) = 0, && \boldsymbol{r} \text{ on } \Gamma\,. && (10.3)
\end{aligned}$$

To progress further, we represent the partial derivatives as differential quotients with notations given in Fig. 10.1:

$$\begin{aligned}
& \partial_t\phi_{out}(\boldsymbol{r}) - \partial_t\phi_{in}(\boldsymbol{r}) = 0 && (10.4) \\
\Leftrightarrow \quad & \lim_{\epsilon\to 0}\left(\frac{\phi(\boldsymbol{r}^+ + \boldsymbol{t}\epsilon) - \phi(\boldsymbol{r}^+)}{\epsilon} - \frac{\phi(\boldsymbol{r}^- + \boldsymbol{t}\epsilon) - \phi(\boldsymbol{r}^-)}{\epsilon}\right) = 0 && (10.5) \\
\Leftrightarrow \quad & \lim_{\epsilon\to 0}\left(\phi(\boldsymbol{r}^+) - \phi(\boldsymbol{r}^-)\right) = \lim_{\epsilon\to 0}\left(\phi(\boldsymbol{r}^+ + \boldsymbol{t}\epsilon) - \phi(\boldsymbol{r}^- + \boldsymbol{t}\epsilon)\right) && (10.6)
\end{aligned}$$

Eq. (10.6) is valid not only for the tangential direction $\boldsymbol{t}$, but for the whole tangential plane. Further, it is valid for every point on Γ, especially for $(\boldsymbol{r}^\pm + \boldsymbol{t}\epsilon)$. Therefore, we conclude that the difference of the potential itself is constant on Γ.

$$\phi(\boldsymbol{x}^+) - \phi(\boldsymbol{x}^-) \quad = \quad \text{const}\,, \quad \text{for all } \boldsymbol{x} \text{ on } \Gamma\,. \tag{10.7}$$

Setting this gauge constant to zero finally yields the continuity of the electrostatic potential on Γ.

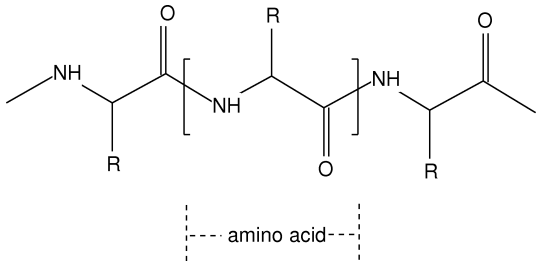


Fig. 10.2: Proteins consist of many amino acids which are connected by the so called peptide bond. The amino acids differ by their side chain R.

10.3 Titration of amino acids

As discussed in Section 3, proteins consist of an arbitrary sequence of natural amino acids. The situation is sketched in Fig. 10.2, where R represents the side chain which specifies the considered amino acid. Some of these side chains are of acidic or basic nature, i.e., they possess functional groups such as a hydroxyl or an amid group, which protonate or deprotonate depending on the solution's pH.

In the present work, we learned that it is of great importance to correctly model the biomolecular system, in particular when we try to predict electrostatic quantities. For instance, a change in the solution's pH can change the charge state of the protein from a negative to a positive total charge. The pH at which a molecule carries no net electrical charge is denoted its *isoelectric point* (pI). The overall behavior of a protein solution highly depends on its characteristic pI value, for it affects the solubility of a molecule at a given pH and finally causes precipitation.

This means that the protonation state as well as the pI value of a protein are important quantities which should be calculated before a further postprocessing. A tool to predict the energetic preferable protonation state is therefore desirable in a biochemical software package such as BALL.

Besides the solvation energy, the protonation state constitutes an indirect measure of the electrostatics present in a biomolecular system. Having this in mind, the study of protonation states is further a means to assess new electrostatic theories, such as the nonlocal electrostatics which is analyzed in the first part of the present work.

Thus, we implemented a *titration simulation program* which numerically estimates the energetically optimal protonation state of all protonable side chains and predicts the protonation state of the protein. The program is integrated in the BALL library as an additional postprocessing tool.

The dissociation constant K_a is a means to estimate the protonation state of acids. In the following sections, we give a brief introduction to the concepts behind biomolecular pK_a values and titration states. First, in Section 10.3.1 we introduce the (measurable) pK_a of an isolated amino acid. In the subsequent section, we transfer this concept to an accumulation of amino acids, i.e., we discuss the corrections made in the case of a protein.

More background information and details on the implementation can be obtained from most biochemistry or biophysical textbooks and from the original articles on pK_a evaluation [11, 111, 150, 151].

10.3.1 Dissociation constant and the pK_a value of an amino acid

Let us consider the equilibrium reactions of a small acid

$$HA \rightleftharpoons H^+ + A^- \,, \tag{10.8}$$

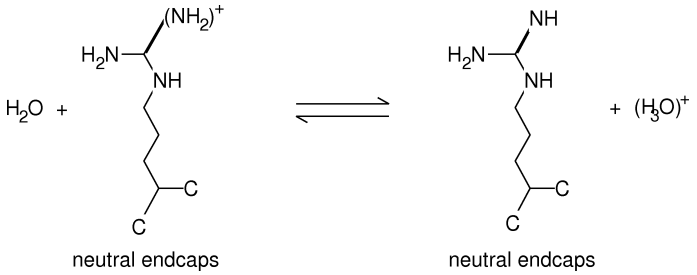


Fig. 10.3: Isolated amino acids are capped with electrostatically neutral ends on both sides. The characteristic side chain of arginine has a protonable amid group. The figure shows the equilibrium reaction of the capped arginine. In dependence of the solvent's pH the equilibrium state is shifted towards the educts (unprotonated) or products (protonated).

or

$$H_2A^+ \rightleftharpoons H^+ + HA\,. \tag{10.9}$$

The latter reaction is shown in Fig. 10.3 for the isolated amino acid arginine with its amid group and neutral endcaps. The dissociation constant K_a of the acid characterized in the following by index a is defined as the equilibrium constant of the reaction given in Eq. (10.9)

$$K_a = e^{-\triangle_{diss}G/(k_{\mathrm{B}}T)}\,,$$

where $\triangle G_{diss} = G_{H^+} + G_{A^-} - G_{HA}$ is the corresponding free energy difference. Here, K_a is the product of the activities of the reactants

$$K_a = \frac{a_H^+\, a_A^-}{a_{HA}}\,. \tag{10.10}$$

Under ideal conditions, activities can be replaced with concentrations to give

$$K_a \approx \frac{c_H^+\, c_A^-}{c_{HA}}\,. \tag{10.11}$$

The dissociation constant can be related to the free energies of the starting materials and the products by

$$-\,k_{\mathrm{B}}\,T \ln K_a = \triangle_{diss}G = G_{H^+} + G_{A^-} - G_{HA}\,. \tag{10.12}$$

Using base-10 instead of the natural logarithm for measuring pH, the pK_a is finally defined as

$$\mathrm{p}K_a = -\log K_a = -\frac{\ln K_a}{\ln 10} = \frac{\triangle_{diss}G}{k_{\mathrm{B}}\,T \ln 10}\,. \tag{10.13}$$

Often, it is of interest to know the energy difference $\triangle_{prot}G$ of the protonated and the unprotonated acid in dependence of the H^+ concentration, i.e., as a function of the solvent's pH, because this reaction takes place when the system equilibrates. Assuming that the activities are replaceable by the concentrations, $\triangle_{prot}G$ equals the logarithm of the ratio between the unprotontated and the protonated acid. By re-arranging Eqs. (10.12) and (10.13) we obtain:

$$\begin{aligned} \triangle_{prot}G &:= G_{HA} - G_{A^-} \approx -k_{\mathrm{B}}\,T \ln \frac{c_{HA}}{c_{A^-}} \\ &= \ln 10\, k_{\mathrm{B}}\,T(\mathrm{p}H - \mathrm{p}K_a)\,. \end{aligned} \tag{10.14}$$

Tab. 10.1: Standard pK_a values of common titratable amino acids.

Amino acid	standard pK_a
Arginine	13.0
Aspartic acid	4.0
Cysteine	8.7
C-terminus	3.8
Glutamic acid	4.4
Histidine	6.3
Lysine	10.4
N-terminus	8.0
Tyrosine	9.6

Eq. (10.14) describes the fact that, when increasing the pH of a solution, the acid will turn from the protonated into the unprotonated state (Henderson-Hasselbalch equation). For $\triangle_{prot}G = 0$, i.e., when the concentrations of unprotonated and protonated acids are equal in the thermodynamic equilibrium, we see with Eq. (10.14) that in this case the pH of the solution equals the pK_a value of the acid.

The pK_a value of the isolated amino acid as, for example, shown for arginine in Fig. 10.3 is measurable in so called *titration* experiments. Tab. 10.1 lists some experimental pK_a values of natural amino acids [111]. As we will see in the next section, these *standard* pK_a values provide the basis for calculating pK_a values of the corresponding amino acid in a protein environment.

10.3.2 pK_a values in proteins

From the experimental view, it is difficult to determine the pK_a values of the amino acids within a protein, because this value highly depends on the chemical surrounding of the protonable amino acids. Further, this means that, even if it had been possible to measure individual dissociation constants of the protonable side chains within a protein environment, these values would not be universal, but only valid for the protein in consideration.

Thus, the idea is to start with the pK_a values of the isolated amino acid and mimic a thermodynamic cycle where the isolated amino acid is transferred into the protein environment. This, then, determines the pK_a value of the considered amino acid in the protein, whereby we assume that all the chemical complexity of protonation, such as bond making and breaking, is correctly represented by the standard pK_a values of the isolated amino acid. The corrections, which we account for by the thermodynamic cycle, are due to the transfer and binding energies required to move the isolated amino acid from water into the protein environment. The thermodynamic cycle is exemplarily depicted for arginine in Fig. 10.4.

1 $H^+(aq) + A^-(aq) \rightarrow HA(aq)$

The corresponding free energy difference, i.e., the energy required or released when changing the protonation state of the amino acid from its unprotonated (0) to its protonated state (1) is determined by Eq. (10.14) with pK_a set to the known standard pK_a, denoted in the following pK_a^{stnd}:

$$\triangle_{prot}G_{(HA(aq):0\rightarrow 1)} = \ln 10\, k_\mathrm{B}\, T(\mathrm{p}H - \mathrm{p}K_a^{stnd})$$

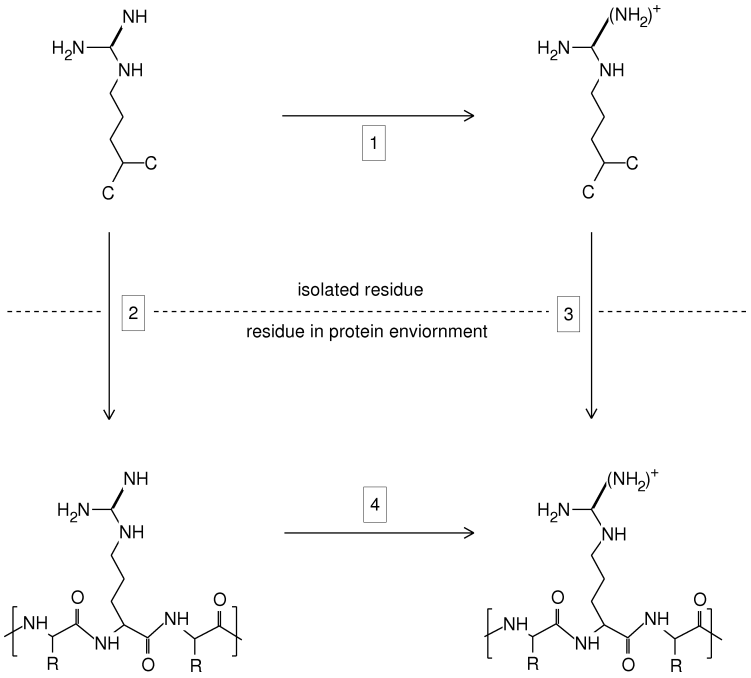


Fig. 10.4: Free energy cycle for the calculation of solvation energies drawn for arginine. The upper row presents the isolated amino acid in its protonated (left) and unprotonated (right) state. The lower row illustrates the same amino acid as a side chain of the protein sequence as implied by brackets on both ends.

2 $A^-(aq) \rightarrow A^-(protein)$
The corresponding (electrostatic) transfer energy $\triangle_{xfer}G_{(A^-)}$ describes the transfer of the isolated unprotonated amino acid into the protein. This energy is approximated by its electrostatic contributions, which can be calculated by solving the Poisson-Boltzmann equation:

$$\triangle_{xfer}G_{(A^-)} = G_{(A^-(protein))} - G_{(A^-(aq))}$$

3 $HA(aq) \rightarrow HA(protein)$
The corresponding (electrostatic) transfer energy $\triangle_{xfer}G_{(HA)}$ describes the transfer of the isolated protonated amino acid into the protein. This energy is approximated by its electrostatic contributions, which can be calculated by solving the Poisson-Boltzmann equation:

$$\triangle_{xfer}G_{(HA)} = G_{(HA(protein))} - G_{(HA(aq))}$$

4 $HA(protein) \rightarrow H^+(aq) + A^-(protein)$
Using the thermodynamic cycle illustrated in Fig. 10.4 we can derive a formula for the corresponding free energy difference when changing the protonation state of the amino acid within the protein from its unprotonated (0) to its protonated state (1):

$$\triangle_{prot}G_{(HA(protein):0\rightarrow 1)} = \triangle_{prot}G_{(HA(aq))} - \triangle_{xfer}G_{(A^-)} + \triangle_{xfer}G_{(HA)}$$

The process we discussed here has not explicitly allowed for changes in the titration state of other groups in the protein during protonation/deprotonation of the acid group of interest. Additionally, it has not explicitly provided for conformational changes in the protein coupled to protonation

and deprotonation. As such, we cannot claim to be computing true pK_a values with this method. Instead, we are calculating so-called intrinsic pK_a values. The p$K_{a,i}^{intr}$ value of amino acid i results from the previously explained thermodynamic cycle:

$$\ln 10\, k_{\mathrm{B}}\, T(\mathrm{p}K_a^{stnd} - \mathrm{p}K_a^{intr}) = \triangle_{xfer} G_{(HA)} - \triangle_{xfer} G_{(A^-)}$$

In contrast to pK_a^{stnd}, the dissociation constants pK_a^{intr} incorporate the effect of the chemical and electrostatic surrounding of the considered amino acid. However, to make such a definition unique we have to define the surrounding, i.e., the titration state of all the other groups in the protein as a certain fixed reference state. For the titration of proteins it is common to assume all titratable amino acid groups to be in their uncharged state as reference.

10.3.3 Free energy difference for given protonation state

With the knowledge of the previous sections, we can calculate the free energy difference, $\Delta G(\boldsymbol{s}^0 \rightarrow \boldsymbol{s}, \mathrm{p}H)$ of an arbitrary state, defined by the state vector $\boldsymbol{s}$, to the reference state $\boldsymbol{s}^0$ at given pH

$$\begin{aligned}\Delta G(\boldsymbol{s}^0 \rightarrow \boldsymbol{s}, \mathrm{p}H) &= (\ln 10)\, k_{\mathrm{B}}\, T \sum_{i=0}^{N} s_i\, (\mathrm{p}H - pK_{a,i}^{intr}) \\ &\quad + \sum_{i=0}^{N} \sum_{j>i}^{N} \left[q_i(s_i)\, q_j(s_j) - q_i(0)\, q_j(0)\right] W_{ij}\,,\end{aligned}$$

$$\text{with} \qquad W_{ij} := E_{ij}(1,1) - E_{ij}(1,0) - E_{ij}(0,1) + E_{ij}(0,0)\,.$$

$E_{ij}(s_i, s_j)$ denotes the interaction energy of the side chain i with side chain j

$$E_{ij}(s_i, s_j) = \sum_l q_{(l,s_j)} \phi_{(i,s_i)}(\boldsymbol{r}_l)\,,$$

where $\boldsymbol{r}_l$ is the position of atom l in the active site j and $\phi_{(i,s_i)}$ is the electrostatic potential of the charge distribution of active residue i in charge state s_i with the dielectric boundary of the protein. The addition/subtraction of the value W_{ij} to/from the free energy ensures that for a given charge state $\boldsymbol{s}$ of the system the electrostatic interaction terms between all atom groups (the titratable and the background groups) are counted only once. The reason for the four terms and their signs is that the product $q_i(s_i)\, q_j(s_j) W_{ij}$ contributes only if both sites are charged, i.e., they differ from the reference state. Thus, the expression for W_{ij} includes the interaction energy E_{ij} with the correct (positive) sign for all combinations of anionic and cationic sites i and j in the considered protonation states $\{s_i, s_j\}$($q_i\, q_j = \pm 1$ if both sites are charged) [111].

10.3.4 Monte Carlo simulation of protein titration

After providing the interaction matrix W_{ij} and the intrinsic pK_a^{intr} values, we have to sample the state space and find the most probable protonation state. Because of the exponential dependence of the number of protonation states on the number of titrating sites, we implemented a Monte Carlo sampling in the way it is proposed in [16, 111] and also explained in Section 8.3.2.2.

The final program consists of two steps: first, electrostatic calculations are executed. This yields the interaction matrix W_{ij} and the pK_a^{intr} vector. Based on these quantities we find the titration curves of all titratable sites by the Monte Carlo simulation. From these curves, the most probable protonation state can be extracted for a given pH. The whole protocol is incorporated in BALL in

order to offer the user to process and work with the correctly protonated biomolecule. Furthermore, this titration program forms the basis for the ion binding experiments discussed in Section 8.3.

As a benchmark and testing example during the implementation, we used the triclinic crystal structure of lysozyme, which is supplied in the examples of the MEAD library [11]. In the following, we compare our results with those generated by MEAD: we prepared exactly the same input information for the electrostatic calculations: first, we adopted the atom configuration, as well as the radius and charge information from MEAD. Second, we adjusted the box dimensions and the molecular surface definitions in the internal FDPB solvers of BALL and MEAD, respectively. Our implementation offers the possibility to read the electrostatic data, which are required for the calculation of the interaction matrix W_{ij} and the pK_{intr} vector from an external electrostatic solver, such as APBS. This gives us the possibility to check the electrostatic calculations separately from the final Monte Carlo procedure and to take advantage of the functionality which is given by APBS.

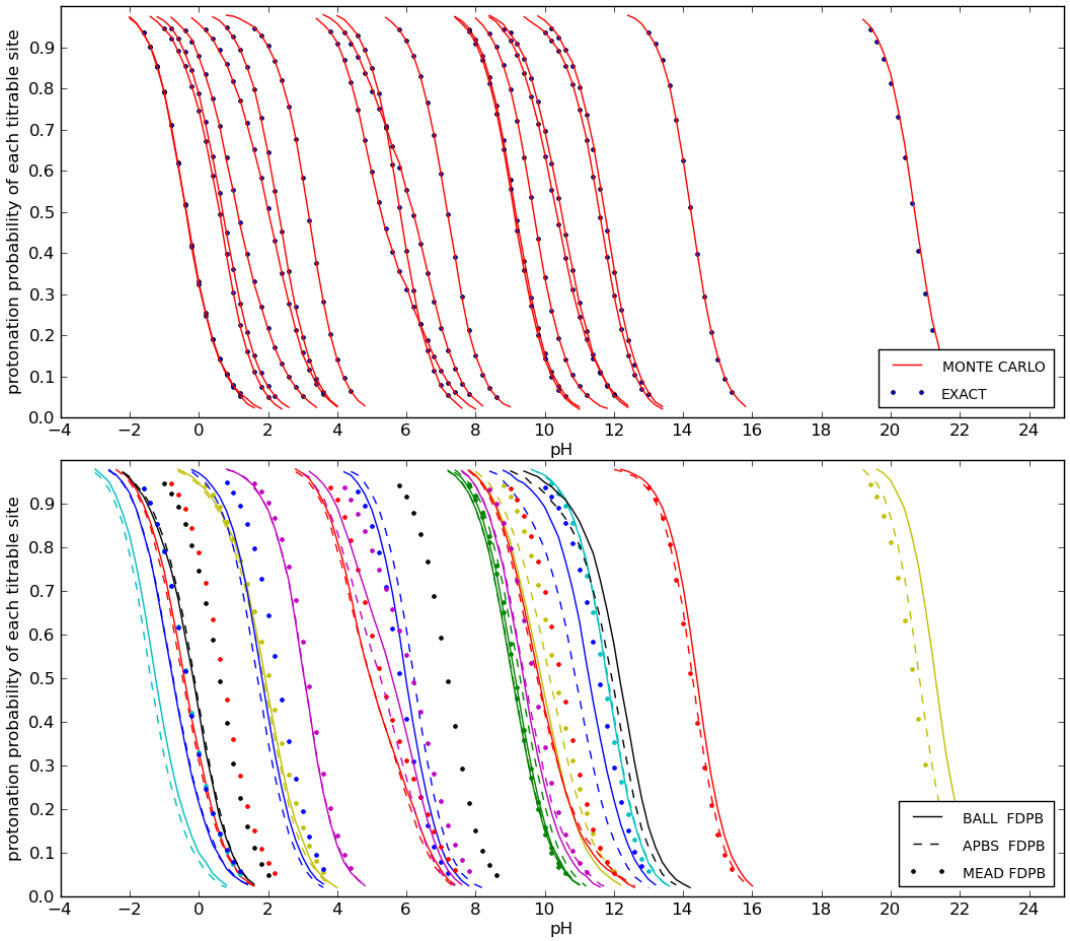


Fig. 10.5: Titration curves of triclinic lysozyme. Upper: data generated by Monte Carlo simulation (lined) and exact sampling of the state space (dotted) with the same interaction matrix and pK_{intr} vector (taken from MEAD). Lower: data generated by the external program MEAD (dotted) and by our Monte Carlo implementation with input generated by (a) BALL finite difference solver (lined) and (b) APBS finite difference solver (dashed). Physical parameters for the electrostatic calculations are $\varepsilon_{in} = 4$, $\varepsilon_{out} = 78.5$, $I = 0$, parameters for the Monte Carlo simulation are $N_{full} = 50000$, $N_{equi} = 50000$.

10.3.4.1 Example: Lysozyme

With the following set of titratable residues {N-Terminus, C-Terminus, LYS, ASP, HIS, TYR, GLU} we obtain a total number $N = 21$ of titratable residues for lysozyme.

Taking the same input in MEAD and in our implementation, we can compare the performance of the Monte Carlo simulation. Fig. 10.5(upper) demonstrates that the Monte Carlo sampling results in the same titration curves as an exact sampling of the state space[1]. Fig. 10.5(lower) shows the results generated by three different finite difference solvers, the BALL internal solver, the MEAD internal solver, and the APBS solver. As the finite difference method sensitively depends on the grid resolution and the molecular surface definition, we started the electrostatic calculations as far as possible with the same options. In summary, the results of all three finite difference solvers are comparable except for one exception appearing in the results of MEAD (black dotted curve with half $pK_a = 7.5$ and black lined and dashed curve with half $pK_a = 12.5$). For both, the APBS and the BALL solver, we used two calculations to estimate the reaction field energies of the protein and the isolated amino acids. As we used MEAD as a black box, such a modification was not possible in MEAD, which explains the differences between the results generated by MEAD and the other two solvers.

10.4 Three possibilities to calculate binding energies

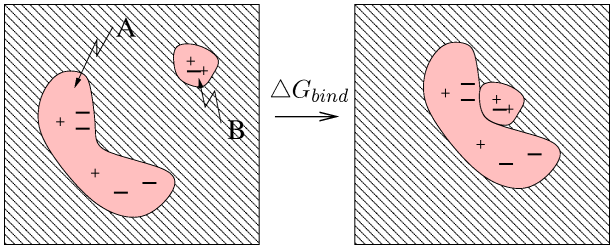


Fig. 10.6: Binding energy by "pure" solvation transitions.

In this section we give an overview on different methods to calculate the energy upon binding of a molecule A and molecule B as sketched in Fig. 10.6. In Section 3.1.4.1, we have introduce the so called solvation energy, here defined as *the energy gain or loss when the dielectric medium in which the charges are immersed is changed*. Two scenarios are illustrated in Fig. 10.7, where the region of the molecule's dielectric response is increased or decreased, respectively.

In Section 3.1.4.1, we learned that the solvation energies can be expressed by the so called reaction field energies, which are based on the reaction field potential. The reason is that in these transitions the charges do not change their position and magnitude. Therefore, their self energy contributions cancel out: assume the potential $\phi(\boldsymbol{r})$ to be the solution of the electrostatic Maxwell's equation and

$$\phi_{mol}(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0\varepsilon_\Omega} \sum_i \frac{q_i}{|\boldsymbol{r} - \boldsymbol{r}_i|}$$

the potential that originates from the external charges within the molecule environment, then, the reaction field potential is defined by

$$\phi_{reac}(\boldsymbol{r}) = \phi(\boldsymbol{r}) - \phi_{mol}(\boldsymbol{r})\,.$$

[1] An exact sampling is no more possible for more than approximately 30 titratable sites.

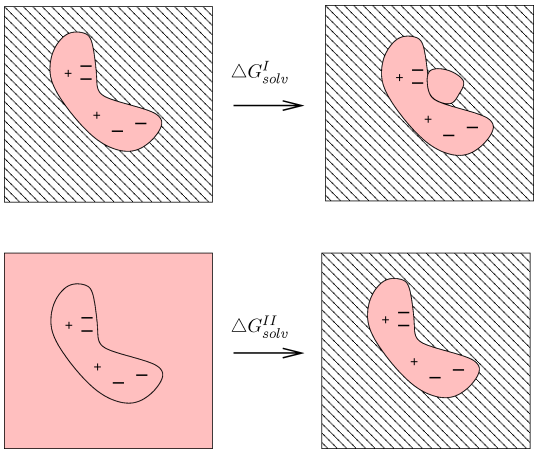


Fig. 10.7: Two different solvation scenarios where the dielectric medium in which the charges are immersed is changed.

When we speak of $(\triangle G^{A}_{solv})$ in the following, we mean the transfer energy of molecule A from the molecule's environment to the solvent environment, i.e., $(\triangle G^{A}_{solv}) = \sum_i q_i \phi_{reac}(\boldsymbol{r}_i)$. When the molecule A is transferred from a dielectric medium to another, where none of them is the molecular surrounding, the solvation energy can be expressed by the difference of the reaction field energies of the two different settings. The problem upon energy calculation is then reduced to the calculation of reaction field energies.

With these definitions, we can express the energy upon binding a molecule B to a molecule A. In the literature several ways are discussed to do this:

- Binding energy can be gained by subtracting the total free energies of the isolated molecules from the one of the complex. This is often used in finite difference calculations, as the grid self energy part cancels because of the subtraction. To completely avoid the grid self energy contribution, one tries to express the thermodynamic cycle in solvation energy cycles as it is done next.

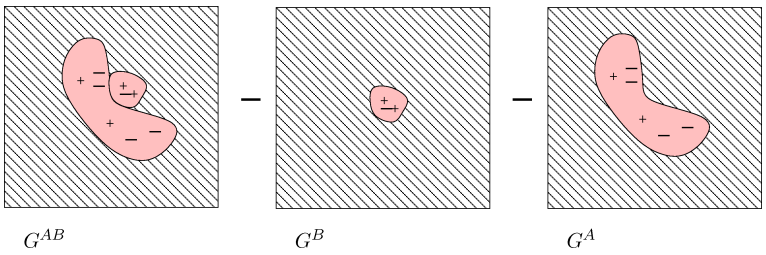


Fig. 10.8: Binding energy by the total free energies.

- An inclusion of solvation energies is given by the following protocol [63]: the change in electrostatic binding energy is separated into three components which are described schematically in Fig. 10.9: (1) the change in solvation energy of the molecule A on binding $(\triangle\triangle G^{A}_{solv})$, (2) the change in solvation energy of the molecule B on binding $(\triangle\triangle G^{B}_{solv})$ and (3) the interaction energy between molecules A and B $(\triangle G^{AB}_{int})$ in the presence of solvent. The binding energy

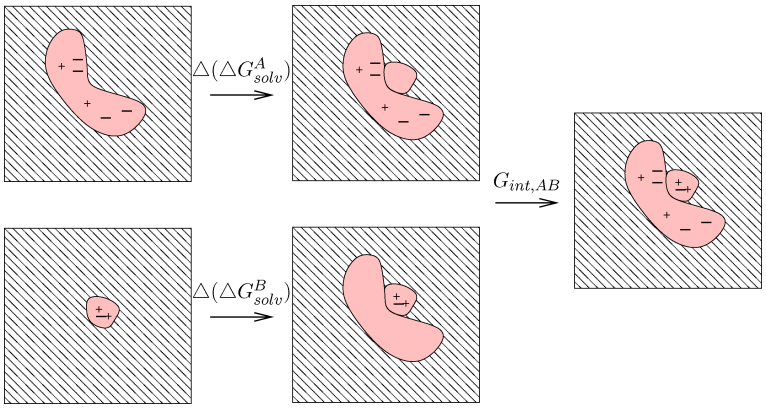


Fig. 10.9: Binding energy calculated by solvation energies and the Coulomb interaction energy of the molecules in the molecule's dielectric environment.

is then given by

$$\triangle G_{bind} = \triangle\triangle G^A_{solv} + \triangle\triangle G^B_{solv} + \triangle G^{AB}_{int} \,.$$

In this formalism, the change in solvation energy corresponds to partial desolvation of the charged molecule A and B on binding by removing the high dielectric solvent from the region occupied by the other molecule on binding and replacing it with the low dielectric medium of the molecules, ε_Ω. This process corresponds to the loss of charge-solvent interaction energy of the molecule on binding. The contributions of A and B can be considered as being independent of one another. In the second step, charges are then transferred to the low dielectric cavity created and the solvent screened interaction energy between molecules is given by

$$G^{AB}_{int} = \sum_{i\in B} q_i\phi_i(A) = \sum_{i\in A} q_i\phi_i(B)\,,$$

where ϕ_i is the potential generated by the charged molecule at the location of a charge q_i in the low dielectric cavity. The interaction energy, G^{AB}_{int} can be calculated either by considering the solvent screened potential generated by the charges on the molecule A at the molecule B, $\phi_i(A)$, or alternatively by that of the charged molecule B at the molecule A, $\phi_i(B)$. Either route should yield an identical value for the interaction energy, subject to very small numerical errors in the finite difference calculation. This overall description of the electrostatic contribution to molecular association gives energies that are easily interpreted as the gain in solvent screened interaction energy at the expense of the loss of solute-solvent interaction energy. This processing is proposed in [63] and is basically used in [111].

- Another way to calculate the binding energy first determines the solvation energy contribution to the binding. This is given by

$$G^{solv}_{bind} = \triangle G^{AB}_{solv} - \triangle G^B_{solv} - \triangle G^B_{solv}\,.$$

G^{solv}_{bind} comprises all the effects of the change in boundary. As already said before, the solvation energies can be expressed by the reaction field energies, which only capture the change in the molecule's surface. This means that G^{solv}_{bind} lacks in considering the intermolecular Coulomb interaction energies between molecule A and B in the complex AB immersed in the uniform molecular dielectric response defined by ε_Ω. This interaction energy can be analytically

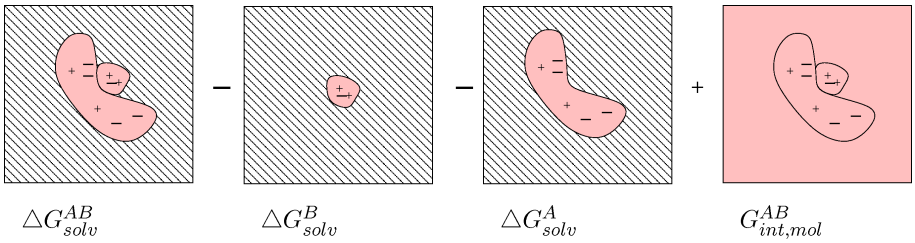


Fig. 10.10: Binding energy by an explicit calculation of the transfer energies.

calculated

$$G^{AB}_{int,mol} = \sum_{i \in B} q_i \phi_{i,mol}(A) = \sum_{i \in A} q_i \phi_{i,mol}(B)\,.$$

To complete the binding free energy cycle, we need to add these contributions to G^{solv}_{bind} to obtain the total electrostatic contribution to the binding free energy

$$G_{bind} = G^{solv}_{bind} + G^{AB}_{int,mol}\,.$$

This processing is proposed in the tutorial of APBS [8].

Bibliography

[1] Johan Åqvist. Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *Journal of Physical Chemicstry*, 94:8021, 1990.

[2] Ramu Anandakrishnan and Alexey Onufriev. Analysis of Basic Clustering Algorithms for Numerical Estimation of Statistical Averages in Biomolecules. *Journal of Computational Biology*, 15(2):165–184, 2008.

[3] J. Antosiewicz, A.J. McCammon, and M.K. Gilson. Prediction of pH-dependent Properties of Proteins. *Journal of Molecular Biology*, 238(3):415–436, 1994.

[4] G. Arfken. *Mathematical Methods for Physicists*. Academic Press, Orlando, FL, 3rd edition, 1985.

[5] Frances H. Arnold and Ji-Hu Zhang. Metal-mediated protein stabilization. *Trends in Biotechnology*, 12(5):189–192, 1994.

[6] Phil Attard, Dongqing Wei, and G.N. Patey. Critical comments on the nonlocal dielectric function employed in recent theories of the hydration force. *Chemical Physics Letters*, 172(1):69–72, 1990.

[7] C. Azuara, H. Orland, M. Bon, P. Koehl, and M. Delarue. Incorporating Dipolar Solvents with Variable Density in Poisson-Boltzmann Electrostatics. *Biophysical Journal*, 95(12):5587–5605, 2008.

[8] N. Baker, A.J. McCammon, and M. Holst. APBS - Adaptive Poisson-Boltzmann Solver. http://www.poissonboltzmann.org/apbs/.

[9] Nathan A. Baker, David Sept, Simpson Joseph, Michael J. Holst, and J. Andrew Mccammon. Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18):10037–10041, 2001.

[10] R. Balakrishnan, N. Ramasubbu, K. I. Varughese, and R. Parthasarathy. Crystal structures of the copper and nickel complexes of RNase A: Metal-induced interprotein interactions and identification of a noval copper binding motif. *Proceedings of the National Academy of Sciences*, 94(18):9620–9625, 1997.

[11] D. Bashford. Macroscopic electrostatic models for protonation states in proteins. *Frontiers in Bioscience*, 9:1082–1099, 2004.

[12] D. Bashford and M. Karplus. pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry*, 29(44):10219–10225, 1990.

[13] Timm Baumeister. *Modellierung der freien Solvatationsenthalpie und Anwendung als Scoringfunktion*. Master's thesis, Institut für Biophysik, Humbolt Universität zu Berlin, Berlin, 2004.

[14] B. Ben-Yaakov, D. Andelman, D. Harries, and R. Podgornik. Beyond standard Poisson-Boltzmann theory: ion-specific interactions in aqueous solutions. *Journal of Physics Condensed Matter*, 21(42):4106–4111, 2009.

[15] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, and P.E. Bourne. The Protein Data Bank. *Nucleic Acids Research*, 28(1):235–242, 2000.

[16] P. Beroza, D. R. Fredkin, M. Y. Okamura, and G. Feher. Protonation of interacting residues in a protein by a Monte Carlo method: application to lysozyme and the photosynthetic reaction center of Rhodobacter sphaeroides. *Proceedings of the National Academy of Sciences*, 88(13):5804–5808, 1991.

[17] Bieniasz, P.D. Intrinsic immunity: a front-line defense against viral attack. *Nature Immunololy*, 5(11):1109–1115, 2004.

[18] Victor A Bloomfield. DNA condensation. *Current Opinion in Structural Biology*, 6(3):334–341, 1996.

[19] C.J.F. Boettcher. *Theory Of Electric Polarization, Volume I: Dielectrics In Static Fields*. Elsevier, 2nd edition, 1973.

[20] Martin Chaplin. *Water*. London South Bank University. http://www1.lsbu.ac.uk/water/chaplin.html.

[21] H.-L. Cheng and X. Shi. Quality mesh generation for molecular skin surfaces using restricted union of balls. *Computational Geometry*, 42(3):196–206, 2009.

[22] H.L. Cheng and H. Edelsbrunner. Skin.exe. http://biogeometry.duke.edu/software/skin/index.html.

[23] A.R. Chipperfield and D.M. Taylor. Binding of Plutonium to Glycoproteins in Vitro. *Radiation Research*, 43(2):393–402, 1970.

[24] P. Cignoni. Meshlab. http://meshlab.sourceforge.net.

[25] E.J. Cohn and J.T. Edsall. *Proteins, Amino Acids, and Peptides*. Reinhold Publishing Corp., New York, 2nd edition, 1943.

[26] D.J. Cram. The design of molecular hosts, guests, and their complexes. *Science*, 240(4853):760–767, 1988.

[27] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Functional and Variational Methods*, volume 2. Springer-Verlag, 1990.

[28] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Physical Origins and Classical Methods*, volume 1. Springer-Verlag, 1990.

[29] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Spectral Theory and Applications*, volume 3. Springer-Verlag, 1990.

[30] M.E. Davis and J.A. McCammon. Electrostatics in Biomolecular Structure and Dynamics. *Chem. Rev.*, 80:509–521, 1990.

[31] M.E. Davis and J.A. McCammon. Dielectric boundary smoothing in finite difference solutions of the poisson equation: An approach to improve accuracy and convergence. *Journal of Computational Chemistry*, 12(7):909–912, 1991.

[32] P. Debye and E. Hückel. Zur Theorie der Elektrolyte. I. Gefrierpunkts–Erniedrigung und Verwandte Erscheinungen. *Phys. Z.*, 24:185, 1923.

[33] Revaz R. Dogonadze, Erika Kálmán, Alexei A. Kornyshev, and Jens Ulstrup. *The Chemical Physics of Solvation, Part A: Theory of Solvation.* Elsevier Science Ltd., 1985.

[34] Revaz R. Dogonadze, Erika Kálmán, Alexei A. Kornyshev, and Jens Ulstrup. *The Chemical Physics of Solvation, Part C.* Elsevier Science Ltd., 1985.

[35] R.R. Dogonadze, A.A. Kornyshev, and A.M. Kuznetsov. Phenomenological description of polar media on the basis of an effective hamiltonian. *Teor. Mat. Fiz., 15:127*, 1973.

[36] Y. Duan and P.A. Kollman. Pathways to a Protein Folding Intermediate Observed in a 1-Microsecond Simulation in Aqueous Solution. *Science*, 282(5389):740–744, 1998.

[37] D. Eisenberg and A.D. McLachlan. Solvation energy in protein folding and binding. *Nature*, 319(6050):199–203, 1986.

[38] U.M. Elofsson, M.A. Paulsson, and T. Arnebrant. Adsorption of β-Lactoglobulin A and B in Relation to Self-Association: Effect of Concentration and pH. *Langmuir*, 13(6):1695–1700, 1997.

[39] C. Fasel. *A boundary integral formulation for nonlocal electrostatics.* Ph.D. thesis, University of Saarland, 2009.

[40] C. Fasel, S. Rjasanow, and O. Steinbach. A boundary integral formulation for nonlocal electrostatics. In *Numerical Mathematics and Advanced Applications - Proceedings of ENUMATH 2007*, pages 117–124. Springer, 2008.

[41] W. Ferenc and A. Walków-Dziewulska. Complexes of Heavy Lanthanides and Yttrium with 3,4-dimethoxybenzoic Acid. *Journal of Thermal Analysis and Calorimetry*, 61(3):923–933, 2000.

[42] N. Fogh-Andersen, P.J. Bjerrum, and O. Siggaard-Andersen. Ionic Binding, Net Charge, and Donnan Effect of Human Serum Albumin as a Function of pH. *Clinical chemistry*, 39(1):48–52, 1993.

[43] E. Gallicchio, L.Y. Zhang, and R.M. Levy. The SGB/NP hydration free energy model based on the surface generalized born solvent reaction field and novel nonpolar hydration free energy estimators. *Journal of Computational Chemistry*, 23(5):517–529, 2002.

[44] Michael Garland and Paul S. Heckbert. Surface simplification using quadric error metrics. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '97, pages 209–216. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1997.

[45] John C. Gordon, Jonathan B. Myers, Timothy Folta, Valia Shoja, Lenwood S. Heath, and Alexey Onufriev. H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic acids research*, 33(Web Server issue):368–371, 2005.

[46] A. Yu. Grosberg, T. T. Nguyen, and B. I. Shklovskii. Colloquium: The physics of charge inversion in chemical and biological systems. *Review of Modern Physics*, 74(2):329–345, 2002.

[47] Bertrand Guillot. A reappraisal of what we have learnt during three decades of computer simulations on water. *Journal of Molecular Liquids*, 101(1-3):219–260, 2002.

[48] T.A. Halgren. Merck Molecular Force Field I-V. *Journal of Computational Chemistry*, 17:490–640, 1996.

[49] T.M. Handel, S.A. Williams, and W.F. DeGrado. Metal ion-dependent modulation of the dynamics of a designed protein. *Science*, 261(5123):879–885, 1993.

[50] J.B. Hasted. *Aqueous Dielectrics.* Chapman and Hall, London, 1973.

[51] M. Heyden, J. Sun, G. Mathias, H. Forbert, M. Havenith, and D. Marx. Dissecting the THz spectrum of liquid water from first principles via correlations in time and space. *Proceedings of the National Academy of Sciences*, 107(27):12068–12073, 2010.

[52] A. Hildebrandt. *Biomolecules in a structured solvent.* Rhombos-Verlag, 2005.

[53] A. Hildebrandt, R. Blossey, S. Rjasanow, O. Kohlbacher, and H.-P. Lenhof. Novel Formulation of Nonlocal Electrostatics. *Phys. Rev. Lett.*, 93(10):108104, 2004.

[54] Francis Hill. *Computer Graphics.* Macmillan USA, 1th edition, 1990.

[55] Lars Hoermander. *The analysis of linear partial differential operators I.* Springer-Verlag, 2nd edition, 1990.

[56] C. Holm, P. Kékicheff, and R. Podgornik. *Proceedings of the NATO advanced research workshop: Electrostatic effects in soft matter and biophysics: Cell model and Poisson-Boltzmann theory: a brief introduction.* NATO science series, 1th edition, 2000.

[57] R.H Holm, P. Kennepohl, and E.I. Solomon. Structural and functional aspects of metal sites in biology. *Chemical Reviews*, 96(7):2239–2314, 1996.

[58] B. et al. Honig. Delphi: A Finite Difference Poisson-Boltzmann Solver.

[59] J.C. Inkson. Dielectric response in semiconductors. *Journal of Physical Chemistry*, 7:1571–1581, 1974.

[60] Jacob Israelachvili. *Intermolecular & Surface Forces.* Academic Press, Harcourt Brace & Company, Publishers, 2nd edition, 1998.

[61] John David Jackson. *Classical Electrodynamics.* John Wiley & Sons, Inc, 3rd edition, 1998.

[62] R.M Jackson and M.J.E Sternberg. Application of scaled particle theory to model the hydrophobic effect: implications for molecular association and protein stability. *Protein Eng.*, 7:371–383, 1994.

[63] R.M Jackson and M.J.E Sternberg. A continuum model for Protein–Protein interactions: Application to the docking problem. *J. Mol. Biol.*, 250:258–275, 1995.

[64] Rainer Jaenicke. Struktur und Funktion von Proteinen. *Radiation and Environmental Biophysics*, 9:299–314, 1973.

[65] Lorin M. Jakubek, Spiro Marangoudakis, Jesica Raingo, Xinyuan Liu, Diane Lipscombe, and Robert H. Hurt. The inhibition of neuronal calcium ion channels by trace levels of yttrium released from carbon nanotubes. *Biomaterials*, 30(31):6351–6357, 2009.

[66] C. T. Kelley. *Iterative Methods for Linear and Nonlinear Equations.* Society for Industrial and Applied Mathematics, Philadelphia, 1995.

[67] A. Kerzmann, D. Neumann, and O. Kohlbacher. SLICK - Scoring and Energy Functions for Protein-Carbohydrate Interactions. *Journal of Chemical Information and Modeling*, 46(4):1635–1642, 2006.

[68] Tönu Kesvatera, Bo Jönsson, Eva Thulin, and Sara Linse. Measurement and Modelling of Sequence-specific pKa Values of Lysine Residues in Calbindin D9k. *Journal of Molecular Biology*, 259(4):828–839, 1996.

[69] R. Knippers. *Molekulare Genetik*. Georg Thieme Verlag, 8th edition, 1998.

[70] O. Kohlbacher H. P. Lenhof et al. BALL – Biochemical Algorithms Library. http://www.bioinf.uni-sb.de/OK/BALL.

[71] Y. Kong and J.W. Ponder. Calculation of the reaction field due to off-center point multipoles. *Journal of Chemical Physics*, 107(2):481–493, 1997.

[72] A.A. Kornyshev, A.I. Rubinstein, and M.A. Vorotyntsev. Model nonlocal electrostatics: I. *J. Phys. C: Solid State Phys.*, 11:3307–3321, 1978.

[73] Kuroishi, A. and Bozek, K. and Shioda, T. and Nakayama, E.E. A single amino acid substitution of the human immunodeficiency virus type 1 capsid protein affects viral sensitivity to TRIM5α. *Retrovirology*, 7(58), 2010.

[74] L.D. Landau and E.M. Lifshitz. *Course of Theoretical Physics, Vol. 8: Electrodynamics of Continuous Media.* Translated from the Russian. Pergamon, Oxford, 1984.

[75] L.D. Landau and E.M. Lifshitz. *Course of Theoretical Physics, Vol. 5: Statistische Physik, Teil 1.* Translated from the Russian. Akademie Verlag Berlin, 1987.

[76] C. Lang and N. Pucker. *Mathematische Methoden in der Physik*. Spektrum, Akademischer Verlag, 1st edition, 1998.

[77] A. Leach. *Molecular Modelling: Principles and Applications.* Pearson Education, 2001.

[78] M.S. Lee, F.R. Salsbury Jr., and C.L. Brooks III. Novel generalized Born methods. *J. Chem. Phys.*, 116(24):10606–10614, 2002.

[79] Z. Li. *The Immersed Interface Method – A Numerical Approach for Partial Differential Equations with Interfaces.* Ph.D. thesis, University of Washington, 1994.

[80] P. Lindqvist-Reis, K. Lamble, S. Pattanaik, I. Persson, and Sandström. Hydration of the Yttrium(III) Ion in Aqueous Solution. An X-ray Diffraction and XAFS Structural Study. *The Journal of Physical Chemistry B*, 104(2):402–408, 2000.

[81] C. Long and B. Mouat. The Binding of Calcium Ions by Erythrocytes and 'Ghost'-Cell Membranes. *Biochemical Journal*, 123:829–836, 1971.

[82] William E. Lorensen and Harvey E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. In *SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 163–169. ACM, New York, NY, USA, 1987.

[83] B.Z. Lu, Y.C. Zhou, M.J. Holst, and J.A. McCammon. Recent Progress in Numerical Methods for the Poisson–Boltzmann Equation in Biophysical Applications. *Communications in Computational Physics*, 3(5):973–1009, 2008.

[84] A.D. Mackerell. Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, 25(13):1584–1604, 2004.

[85] J.D. Madura, J.M. Briggs, R.C. Wade, M.E. Davis, B.A. Luty, A. Ilin, J. Antosiewicz, M.K. Gilson, B. Bagheri, L.R. Scott, and J.A. McCammon. Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics program. *Computer Physics Communications*, 91:57–95, 1995.

[86] A. C. Maggs and R. Everaers. Simulating Nanoscale Dielectric Response. *Phys. Rev. Lett.*, 96(23):230603, 2006.

[87] Y. Marcus. *Ion Solvation*. Wiley, New York, 1985.

[88] M. Marquart, J. Walter, J. Deisenhofer, W. Bode, and R. Huber. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta Crystallographica Section B*, 39(4):480–490, 1983.

[89] D.R. Mattisson and F.A.C. Mattison. Sex Differences in Drug Development. *Blickpunkt der Mann*, 6:21–25, 2008.

[90] William McLean. *Strongly Elliptic Systems and Boundary Integral Equations*. Cambridge University Press, Cambridge CB2 2RU, UK, 1st edition, 2000.

[91] Merck and Inc. Co. MMFF94 validation suite. http://ftp.ccl.net/cca/data/MMFF94/index.shtml.

[92] E.L. Mertz and L.I. Krishtalik. Low dielectric response in enzyme active site. *Proceedings of the National Academy of Sciences*, 97(5):2081–2086, 2000.

[93] A.C. Metaxas and R.J. Meredith. *Industrial microwave heating*. Peregrinus, London, 1th edition, 1983.

[94] W.G. Mumme and A.D. Wadsley. The structure of orthorhombic Y_2TiO_5, an example of mixed seven- and fivefold coordination. *Acta Crystallographica Section B*, 24(10):1327–1333, 1968.

[95] D.A. Nachshen. Selectivity of the Ca binding site in synaptosome Ca channels. Inhibition of Ca influx by multivalent metal cations. *Journal of Cell Biology*, 83(6):941–967, 1984.

[96] T. Nicolai, M. Pouzot, D. Durand, M. Weijers, and R.W. Visschers. Iso-scattering points during heat-induced aggregation and gelation of globular proteins indicating micro-phase separation. *Europhysical Letters*, 73(2):299–305, 2006.

[97] Wolfgang Nolting. *Grundkurs: Theoretische Physik 3: Elektrostatik*. Springer, Berlin, 5th edition, 2001.

[98] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham, S.E. DeBolt, D.M. Ferguson, G.L. Seibel, and P.A. Kollman. AMBER, a Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comp. Phys. Let.*, 1995.

[99] J. Ray and G.S. Manning. An attractive force between two rodlike polyions mediated by the sharing of condensed counterions. *Langmuir*, 10(7):2450–2461, 1994.

[100] Lynne Regan. Protein design: novel metal-binding sites. *Trends in Biochemical Sciences*, 20(7):280–285, 1995.

[101] W. Rocchia, S. Sridharan, A. Nicholls, E. Alexov, A. Chiabrera, and B. Honig. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. *Journal of Computational Chemistry*, 23(1):128–137, 2002.

[102] Jörg Rottler and B. Krayenhoff. Numerical studies of nonlocal electrostatic effects on the sub-nanoscale. *Journal of Physics: Condensed Matter*, 21(25):255901, 2009.

[103] I. Rouzina and V.A. Bloomfield. Macroion Attraction Due to Electrostatic Correlation between Screening Counterions. 1. Mobile Surface-Adsorbed Ions and Diffuse Ion Cloud. *The Journal of Physical Chemistry*, 100(23):9977–9989, 1996.

[104] V. Rutka. *Immersed Interface Methods for Elliptic Boundary Value Problems*. Ph.D. thesis, TU Kaiserslautern, 2005.

[105] V. Rutka. A staggered grid based explicit jump immersed interface method for two-dimensional Stokes flows. *Int. J. Numer. Meth. Fluids*, 57:1527–1543, 2008.

[106] V. Rutka and A. Wiegmann. Explicit Jump Immersed Interface Method: Documentation for 2D Poisson Code. http://service.ifam.uni-hannover.de/∼rutka/forschung/EjiimEx/DOC/documentation/.

[107] V. Rutka and A. Wiegmann. Explicit jump immersed interface method for virtual material design of the effective elastic moduli of composite materials. *Numer. Algor.*, 43:309–330, 2006.

[108] R.S. Sandhu. A Thermodynamic Study of Complexation Reaction of Yttrium(III), Lanthanum(III) and Cerium(III) with Tyrosine. *Monatshefte für Chemie*, 108(1):51–55, 1977.

[109] M.F. Sanner, M.J. Olson, and J.-C. Spehner. Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers*, 38(3):305–320, 1996.

[110] G. Scatchard, I.H. Scheinberg, and S.H. Armstrong. Physical Chemistry of Protein Solutions. V. The Combination of Human Serum Albumin with Thiocyanate Ion. *Journal of the American Chemical Society*, 72(1):540–546, 1950.

[111] M. Schaefer, M. Sommer, and M. Karplus. pH-Dependence of Protein Stability: Absolute Electrostatic Free Energy Differences between Conformations. *The Journal of Physical Chemistry B*, 101(9):1663–1683, 1997.

[112] Scott, Ridgway and Boland, Mercedes and Rogale, Kristina and Fernández, Ariel. Continuum equations for dielectric response to macro-molecular assemblies at the nano scale. *Journal of Physics A: Mathematical and General*, 37(41):9791, 2004.

[113] J.A. Sethian and A. Wiegmann. Structural Boundary Design via Level Set and Explicit Jump Immersed Interface Methods. *Journal of Computational Physics*, 163(2):489–528, 2000.

[114] X. Shi and P. Koehl. The Geometry Behind Numerical Solvers of the Poisson-Boltzmann Equation. *Communications in Computational Physics*, 3(5):1032–1050, 2008.

[115] H. Sigel and M.R. Bruce. Coordinating properties of the amide bond. Stability and structure of metal ion complexes of peptides and related ligands. *Chemical Reviews*, 82(4):385–426, 1982.

[116] Doree Sitkoff, Kim A. Sharp, and Barry Honig. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem.*, 98:1978–1988, 1994.

[117] X. Songa. An inhomogeneous model of protein dielectric properties: Intrinsic polarizabilities of amino acids. *Journal of chemical physics*, 116(21):9359–9363, 2002.

[118] A.K. Soper. Orientational correlation function for molecular liquids: The case of liquid water. *J. Chem. Phys.*, 101(8):6888–6901, 1994.

[119] Noel T. Southall, Ken A. Dill, and A.D.J. Haymet. A View of the Hydrophobic Effect. *The Journal of Physical Chemistry B*, 106(3):521–533, 2002.

[120] P.E. Stein, A.G. Leslie, J.T. Finch, and R.W. Carrell. Crystal structure of uncleaved ovalbumin at 1.95 A resolution. *Journal of Molecular Biology*, 222(3):941–959, 1991.

[121] Olaf Steinbach. *Numerische Näherungsverfahren für elliptische Randwertprobleme – Finite Elemente und Randelemente (in German).* Advances in Numerical Mathematics. Teubner Verlag/GWV Fachverlage GmbH, Wiesbaden, 2003.

[122] M. Stremlau, M. Perron, M. Lee, Y. Li, B. Song, H. Javanbakht, F. Diaz-Griffero, D.J. Anderson, W.I. Sundquist, and J. Sodroski. Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5α restriction factor. *Proceedings of the National Academy of Sciences*, 103(14):5514–5519, 2006.

[123] Stremlau, M. and Owens, C.M. and Perron, M.J. and Kiessling, M. and Autissier, P. and Sodroski, J. The cytoplasmic body component TRIM5α restricts HIV-1 infection in Old World monkeys. *Nature*, 427(6977):848–853, 2004.

[124] Godehard Sutmann. *Die nichtlokale dielektrische Funktion von Wasser.* Ph.D. thesis, Forschungszentrum Jülich, Institut für Werkstoffe und Verfahren der Energietechnik, 1999.

[125] J. Swanson, S. Adcock, and J.A. McCammon. Optimized Radii for Poisson-Boltzmann Calculations with the AMBER Force Field. *Journal of Chemical Theory and Computation*, 1(3):484–493, 2005.

[126] H.A. Tajmir-Riahi, R. Ahmad, and M. Naoui. Interaction of calf-thymus DNA with trivalent La, Eu, and Tb ions. Metal ion binding, DNA condensation and structural features. *Journal of Biomolecular Structure and Dynamics*, 10(5):865–877, 1993.

[127] C. Tanford, S. Swanson, and W. Shore. Hydrogen Ion Equilibria of Bovine Serum Albumin. *Journal of the American Chemical Society*, 77(24):6414–6421, 1955.

[128] Tang, C. and Ndassa, Y. and Summers, M.S. Structure of the N-terminal 283-residue fragment of the immature HIV-1 Gag polyprotein. *Nature Structural Biology*, 9(7):537–543, 2002.

[129] P.A. Tishmack, D. Bashford, E. Harms, and R.L. Van Etten. Use of 1H NMR Spectroscopy and Computer Simulations To Analyze Histidine pKa Changes in a Protein Tyrosine Phosphatase: Experimental and Theoretical Determination of Electrostatic Properties in a Small Protein. *Biochemistry*, 36(39):11984–11994, 1997.

[130] D.M. Tully-Smith and H. Reiss. Further Development of Scaled Particle Theory of Rigid Sphere Fluids. *Journal of Chemical Physics*, 53(10):4015–4025, 1970.

[131] E.K.M. Ueda, P.W. Gout, and L. Morganti. Current and prospective applications of metal ion-protein binding. *Journal of Chromatography A*, 988(1):1–23, 2003.

[132] H.H. Uhlig. The Solubilities of Gases and Surface Tension. *J. Phys. Chem.*, 41(9):1215–1226, 1937.

[133] Arthur Robert von Hippel. *Dielectric materials and applications*. Artech House, Boston, 1th edition, 1995.

[134] M.A. Vorotyntsev. Model nonlocal electrostatics: II. Spherical interface. *J. Phys. C: Solid State Phys.*, 11:3323–3331, 1978.

[135] Jason A. Wagoner and Nathan A. Baker. Assessing implicit models for nonpolar mean solvation forces: The importance of dispersion and volume terms. *Proceedings of the National Academy of Sciences*, 103(22):8331–8336, 2006.

[136] Lianghua Wang, Shaochuan Shen, Xingjiao He, Junxian Yun, Kejian Yao, and Shan-Jing Yao. Adsorption and elution behaviors of bovine serum albumin in metal-chelated affinity cryogel beds. *Biochemical Engineering Journal*, 42(3):237–242, 2008.

[137] Kyoko Watanabe and Michael L. Klein. Effective pair potentials and the properties of water. *Chemical Physics*, 131(2-3):157–167, 1989.

[138] S. Weggler, V. Rutka, and A. Hildebrandt. A new numerical method for nonlocal electrostatics in biomolecular simulations. *Journal of Computational Physics*, 229(11):4059–4074, 2010.

[139] R. Clint Whaley, Antoine Petitet, and Jack J. Dongarra. Automated Empirical Optimization of Software and the ATLAS Project. *Parallel Computing*, 27(1–2):3–35, 2001.

[140] A. Wiegmann. *The Explicit–Jump Immersed Interface Method and Interface Problems for Differential Equations*. Ph.D. thesis, University of Washington, 1998.

[141] A. Wiegmann. Fast Poisson, fast Helmholtz and fast linear elastostatic solvers on rectangular parallelepipeds. Technical Report LBNL-43565, Lawrence Berkeley National Laboratory, MS 50A-1148, One Cyclotron Rd, Berkeley CA 94720, 1999.

[142] A. Wiegmann. The Explicit Jump Immersed Interface Method and Integral Formulas. Technical Report LBNL-43566, Lawrence Berkeley National Laboratory, MS 50A-1148, One Cyclotron Rd, Berkeley CA 94720, 1999.

[143] A. Wiegmann and K.P. Bube. The immersed interface method for nonlinear differential equations with discontinuous coefficients and singular sources. *SIAM J. Numer. Anal.*, 35(1):177–200, 1998.

[144] A. Wiegmann and K.P. Bube. The Explicit-Jump Immersed Interface Method: Finite Difference Methods for PDEs with Piecewise Smooth Solutions. *SIAM J. Numer. Anal.*, 37(3):827–862, 2000.

[145] P.A. Williams and A.R. Peacocke. The Binding of Calcium and Yttrium Ions to a Glycoprotein from Bovine Cortical Bone. *Biochemical Journal*, 105:1177–1185, 1967.

[146] T. You and D. Bashford. An analytical algorithm for the rapid determination of the solvent accessibility of points in a three-dimensional lattice around a solute molecule. *Journal of Computational Chemistry*, 16(6):743–757, 1995.

[147] H. Zahn and K. Mella. Nα-Acetyl-methylamide von Lysin, Histidin, Serin, Tyrosin, Cystin und Tryptophan. *Hoppe-Seyler´s Zeitschrift für physiologische Chemie*, 344(1-3):75–82, 1966.

[148] R.J. Zauhar. SMART: a solvent–accessible triangulated surface generator for molecular graphics and boundary element applications. *J. Comput. Aided Mol. Des.*, 9:149–159, 1995.

[149] F. Zhang. Private communication with Dr. Fajun Zhang (Department of physics, Tübingen).

[150] F. Zhang, M.W.A. Skoda, R.M.J. Jacobs, S. Zorn, R.A. Martin, C.M. Martin, G.F. Clark, S. Weggler, A. Hildebrandt, O. Kohlbacher, and F. Schreiber. Reentrant Condensation of Proteins in Solution Induced by Multivalent Counterions. *Physical Review Letters*, 101(14):148101, 2008.

[151] F. Zhang, S. Weggler, J.M. Ziller, L. Ianeselli, B.S. Heck, A. Hildebrandt, O. Kohlbacher, M.W. Skoda, R.M.J. Jacobs, and F. Schreiber. On the Universality of Protein Reentrant Condensation in Solution Induced by Multivalent Counterions. *Proteins: Structure, Function, and Bioinformatics*, 78(16):3450–3457, 2010.

[152] Ruhong Zhou. Free energy landscape of protein folding in water: Explicit vs. implicit solvent. *Proteins: Structure, Function, and Genetics*, 53(2):148–161, 2003.