

A new quasilinear saturation rule for  
tokamak turbulence

Harry George Dudding

Doctor of Philosophy

University of York

Physics

October 2022

## Abstract

The confinement of tokamak plasmas is largely determined by the presence of turbulence, as a result of its associated radial transport of energy and particles. This level of transport is observed to vary with isotope mass, however with an experimental scaling relation in opposition to that suggested by simple theory. A greater understanding of the isotope dependence of confinement is therefore required for the development of operating scenarios in future devices. Integrated plasma simulators can be used to make predictions of confinement, which calculate transport via quasilinear turbulence models. The simplifications employed in these models to enable their required computational efficiency can cause their transport predictions to deviate from those of the more accurate but prohibitively expensive framework of nonlinear gyrokinetics, motivating their continued verification and improvement.

Steep gradients in the plasma density characteristic of the tokamak edge coupled with the non-adiabatic response of the electrons can drive turbulence dominated by the trapped electron mode. Nonlinear gyrokinetic simulations demonstrate that the resulting local transport exhibits isotope scaling reversal, which contemporary quasilinear models are unable to replicate due to the absence of the relevant physics in their description of the turbulence. Through analysis of gyrokinetic spectra this work attributes the physical origin of this discrepancy to the saturation level of the fluctuating electrostatic potential. A new quasilinear rule SAT3 is constructed to extend previous descriptions of turbulent saturation via the incorporation of differing saturation levels depending on the turbulence regime. This enables the isotope scaling reversal of the trapped electron mode to be described quasilinearly for the first time, whilst retaining the more established description of ion temperature gradient driven turbulence in the appropriate parameter space.

The new model is validated against data from the recent JET isotope experiments in H, D and T. SAT3 is seen to perform well in integrated modelling simulations, however further investigation into the relative contributions of the various physical effects present is required.

# Contents

<b>Abstract</b>	<b>1</b>
<b>Contents</b>	<b>2</b>
<b>List of Tables</b>	<b>7</b>
<b>List of Figures</b>	<b>8</b>
<b>Acknowledgements</b>	<b>15</b>
<b>Declaration</b>	<b>16</b>
<b>1 Introduction</b>	<b>17</b>
1.1 Fusion energy . . . . .	17
1.1.1 Fusion power balance . . . . .	19
1.2 Physics of plasmas . . . . .	21
1.2.1 Plasma models . . . . .	21
1.2.2 Adiabatic response . . . . .	24
1.2.3 Quasineutrality . . . . .	24
1.2.4 Particle motion in uniform electric and magnetic fields . . . . .	25
1.3 Plasma confinement . . . . .	29
1.4 Tokamaks . . . . .	31
<b>2 Tokamak confinement</b>	<b>36</b>
2.1 Collisional transport . . . . .	36
2.2 Experimental character of anomalous transport . . . . .	37
2.3 Turbulent transport . . . . .	37
2.3.1 Electron drift waves . . . . .	38
2.3.2 Other drift waves . . . . .	42

2.4	Empirical scaling laws and modes of tokamak operation . . . . .	43
2.5	First principles modelling . . . . .	45
2.5.1	Turbulent transport models . . . . .	46
2.6	The isotope effect . . . . .	48
2.7	Topics addressed in this thesis . . . . .	50
<b>3</b>	<b>Gyrokinetic theory</b>	<b>52</b>
3.1	Introduction . . . . .	52
3.1.1	The ensemble average . . . . .	52
3.1.2	Gyrokinetic ordering assumptions . . . . .	54
3.2	Equations of motion . . . . .	56
3.2.1	Phase coordinates . . . . .	56
3.2.2	Lowest order solution . . . . .	59
3.2.3	First order . . . . .	61
3.2.4	Velocity coordinates . . . . .	63
3.2.5	Single particle motion summary . . . . .	65
3.3	The gyrokinetic equation . . . . .	66
3.3.1	Zeroth order: gyrophase-independence of $F_{s,0}$ . . . . .	67
3.3.2	First order: Maxwellian background and the splitting of $\delta f_{s,1}$ . . . . .	67
3.3.3	Second order: the gyrokinetic equation . . . . .	69
3.3.4	Maxwell's equations . . . . .	69
3.4	Tokamak geometry . . . . .	71
3.4.1	Divergenceless condition . . . . .	71
3.4.2	Axisymmetric flux surfaces . . . . .	72
3.4.3	Toroidal coordinates . . . . .	73
3.4.4	Plasma equilibrium: force balance and Ampère's law . . . . .	75
3.4.5	The local Grad-Shafranov solution in toroidal coordinates . . . . .	77
3.4.6	Plasma shape . . . . .	79
3.4.7	Toroidal coordinate summary . . . . .	81
3.4.8	Global field-aligned coordinates . . . . .	82
3.4.9	The local limit . . . . .	85
3.5	Local gyrokinetics . . . . .	90
3.5.1	Normalisations . . . . .	91
3.6	Turbulent transport in local gyrokinetics . . . . .	91

3.6.1	Flux calculation . . . . .	91
3.6.2	Turbulent flux characteristics . . . . .	96
3.7	Outputs of the local gyrokinetic system . . . . .	97
3.7.1	Nonlinear gyrokinetics . . . . .	97
3.7.2	Linear gyrokinetics . . . . .	100
<b>4</b>	<b>Quasilinear models</b>	<b>102</b>
4.1	Introduction . . . . .	102
4.2	Quasilinear theory . . . . .	102
4.2.1	The weight spectrum, $W_{s,k_y}$ . . . . .	103
4.2.2	The saturated potential spectrum, $\langle  \delta\hat{\phi}_{k_y} ^2 \rangle_{x,\theta,t}$ . . . . .	105
4.3	Fluxes in reduced models . . . . .	106
4.3.1	Flux calculation . . . . .	107
4.4	TGLF and QuaLiKiz . . . . .	108
4.4.1	TGLF . . . . .	108
4.4.2	QuaLiKiz . . . . .	111
4.5	Isotope scaling in quasilinear models . . . . .	114
4.5.1	GyroBohm scaling in local gyrokinetic systems . . . . .	114
4.5.2	Incorrect isotope scaling in quasilinear models . . . . .	114
4.6	Simulation database . . . . .	115
4.7	Model comparison and isotope scaling diagnosis . . . . .	117
4.7.1	Isotope scaling metric $\alpha_A$ . . . . .	117
4.7.2	Flux comparison . . . . .	118
4.7.3	Comparison of flux constituents . . . . .	120
<b>5</b>	<b>Development of a new saturation rule</b>	<b>123</b>
5.1	2D spectrum . . . . .	123
5.2	Saturation equations . . . . .	127
5.2.1	Radial spectral width parameterisation . . . . .	127
5.2.2	Parameter considerations . . . . .	129
5.2.3	Model for the radial spectral width . . . . .	129
5.3	Physical interpretation of saturation rule parameters . . . . .	130
5.4	Approximating the parameters from linear physics quantities . . . . .	132
5.5	Model extensions . . . . .	135
5.5.1	Subdominant modes . . . . .	135

5.5.2	$E \times B$ shear . . . . .	135
5.5.3	Mixed plasmas . . . . .	136
5.5.4	Electron scale . . . . .	136
5.6	Quasilinear approximation functions . . . . .	138
5.7	Summary of new saturation model, SAT3 . . . . .	140
5.8	Results . . . . .	142
5.9	A note on future improvements . . . . .	146
<b>6</b>	<b>Validation against JET isotope experiments</b>	<b>147</b>
6.1	Introduction . . . . .	147
6.2	JETTO . . . . .	148
6.3	JET pulses . . . . .	149
6.3.1	L-mode pulses . . . . .	149
6.3.2	Ohmic pulses . . . . .	151
6.4	JETTO simulations . . . . .	154
6.4.1	L-mode simulations . . . . .	157
6.4.2	Ohmic simulations . . . . .	160
6.5	Summary and future works . . . . .	164
<b>7</b>	<b>Summary and discussion</b>	<b>166</b>
<b>A</b>	<b>Background concepts</b>	<b>172</b>
A.1	Velocity moments of the kinetic equation . . . . .	172
A.1.1	General velocity moment . . . . .	172
A.1.2	The evolution of density and momentum . . . . .	173
A.2	Quasineutrality . . . . .	176
A.3	Phase in Fourier modes . . . . .	176
<b>B</b>	<b>Gyrokinetic results</b>	<b>178</b>
B.1	Time average . . . . .	178
B.2	First order gyromotion contribution . . . . .	178
B.3	Properties of the gyroaverage . . . . .	181
B.4	Velocity coordinate evolution equations . . . . .	182
B.5	Derivation of the gyrokinetic equation . . . . .	183
B.6	The gyrokinetic equation in general velocity coordinates . . . . .	186

<b>C</b>	<b>Differential geometry</b>	<b>190</b>
C.1	Cartesian results . . . . .	190
C.2	General coordinate systems . . . . .	192
C.2.1	Basis vectors . . . . .	192
C.2.2	Vector representations . . . . .	195
C.2.3	Vector operations . . . . .	198
C.2.4	Transforming between systems . . . . .	200
<b>D</b>	<b>Equilibrium geometry</b>	<b>201</b>
D.1	Equilibrium plasma momentum equation . . . . .	201
D.2	Cylindrical geometry . . . . .	202
D.2.1	Axisymmetric divergencelessness . . . . .	202
D.3	Toroidal coordinates . . . . .	203
D.3.1	The potential $G$ in toroidal coordinates . . . . .	204
D.3.2	Magnetic fluxes . . . . .	205
D.3.3	Alternative angular coordinates . . . . .	206
D.3.4	The standard Grad-Shafranov equation . . . . .	209
D.4	Global field-aligned coordinates . . . . .	210
D.4.1	Third coordinate considerations . . . . .	210
D.4.2	Geometric quantities . . . . .	211
D.5	Local field-aligned system . . . . .	213
D.5.1	Geometric quantities . . . . .	213
D.5.2	Terms in the local gyrokinetic equation . . . . .	214
D.5.3	Gyroaverage operators . . . . .	215
<b>E</b>	<b>Developing SAT3</b>	<b>217</b>
E.1	Parseval's theorem . . . . .	217
E.2	Grid-resolution dependence . . . . .	217
E.3	Database convergence study . . . . .	218
E.4	2D spectrum moment integrals . . . . .	220
	<b>Bibliography</b>	<b>223</b>

# List of Tables

4.1	Details of the 43 nonlinear CGYRO cases that form the database. This set was generated to recreate a subset of the points of the density scan seen in figure 4.2, and then expanded out in a number of parameters of interest from the well-studied GA-std case to test the quasilinear models' dependence in these different areas. Note that the labels in the second column correspond with those in figure 5.8c, used for discerning the cases displayed in figures 5.8a, 5.8b, 5.12 and 5.13. . . . .	116
6.1	Experimental details of the 12 Ohmic cases simulated. Note that the measured $Z_{\text{eff}}$ values are unreliable due to the difficulty in its experimental measurement. A subset of the values are hence changed for the simulations of this chapter, as discussed in the text. . . . .	152
6.2	Parameters for the NBI heating used in PENCIL for the L-mode simulations.	155
6.3	Isotope scaling metric values between different shots for the L-mode set with all effects included on the left hand column, and no momentum transport on the right hand column. . . . .	158
6.4	Isotope scaling metric values between different isotopes for the low- and mid-density Ohmic sets. . . . .	161
6.5	Isotope scaling metric values between different isotopes for the two high-density Ohmic sets. . . . .	162

# List of Figures

1.1	Relative velocity-averaged cross sections $\langle\sigma v\rangle$ as a function of temperature for a selection of low $Z$ fusion reactions. Figure adapted from [3], with the data obtained from [4]. . . . .	18
1.2	The data of figure 1.1 now divided by the square of the temperature. Note the appearance of a global maximum for each reaction. . . . .	21
1.3	Particle gyromotion in the presence of a uniform magnetic field and no electric field. Note that the anti-clockwise orientation of the motion corresponds to an electron. . . . .	28
1.4	$E \times B$ motion of an electron (grey) and ion (red) in the presence of constant orthogonal electric and magnetic fields. The dynamics parallel to the magnetic field have been suppressed for simplicity. . . . .	28
1.5	Schematic illustrating the relation between the particle position $\mathbf{r}$ , the gyrocentre $\mathbf{R}$ , the gyroradius $\rho$ and the gyrophase $\varphi$ . . . . .	29
1.6	Exploring the case of attempting plasma confinement using a concentric circular magnetic field. The drifts resulting from the field configuration would cause a separation of space charge between the ions and electrons, generating an $E \times B$ drift which would cause confinement to be lost. . . . .	31
1.7	Schematic of the magnetic field configuration in a tokamak [13]. . . . .	32
1.8	Schematic of flux surfaces present in a tokamak, with coordinate labels for the radial, poloidal and toroidal directions $\{\varrho, \vartheta, \zeta\}$ . Figure adapted from [14]. . .	33
2.1	The mechanism of the electron drift wave, at some initial time $t = 0$ (a) and a quarter of a period later, $t = T_\omega/4$ (b). Note the movement of the wave in the positive $y$ direction between (a) and (b). . . . .	40
2.2	A drift wave now with a phase shift of $\pi/2$ between the density and the potential. The resulting $E \times B$ motion now acts to further perturb the density, resulting in an instability. . . . .	41

3.1	Cartoon of the cylindrical $\{R, \Phi, Z\}$ and toroidal $\{\varrho, \vartheta, \zeta\}$ coordinate systems. Note $\zeta$ is directed out of the page, whereas $\Phi$ is directed into the page. . . . .	73
3.2	Comparison of the quantity $g^{\alpha\alpha}$ obtained from the Mercier-Luc formalism using CGYRO (coloured circles), and that obtained in this section (white stars) for a selection of tokamak equilibria. Provided the other terms in the equation are known, $g^{\alpha\alpha}$ becomes a proxy for agreement between the radial derivatives in the system. For all equilibria the curves overlap everywhere, indicating the equivalence of the two approaches. . . . .	79
3.3	Examples of flux surfaces constructed using the Miller parameterisation with various anchor points marked. Note that the poloidal angle does not necessarily correspond to the parameterisation angle. . . . .	81
3.4	Schematic of the field-aligned system $\{r, \alpha, \theta\}$ , including the alignment of the equilibrium magnetic field $\mathbf{B}$ with the covariant basis vector of the $\theta$ coordinate, $\mathbf{a}_\theta$ . . . . .	83
3.5	Schematic of the local limit domain of the ‘flux tube’, centered around a single equilibrium field line. Note the flattening of the tokamak flux surface is for illustrative purposes only, and the lines have been displayed as straight for simplicity, not as a consequence of flux coordinates. . . . .	86
3.6	(a) The trace of the total ion heat flux in the GA-std case in D, exhibiting the method of time averaging used in this work. (b) The flux component spectrum $Q_{i,k_y}$ of the ion heat flux in D for the GA-std case. . . . .	98
3.7	The 1D saturated potential spectrum for the GA-std case in D, shown with (a) and without (b) the zonal mode. . . . .	99
3.8	Nonlinear weight spectrum for the GA-std case in D. . . . .	99
3.9	Eigenvalue spectra for the linear gyrokinetic system of the GA-std case in D, showing the growth rates $\gamma_{k_y}$ (a) and the real frequencies $\tilde{\omega}_{k_y}$ (b). . . . .	100
3.10	The linear weight spectrum for the ion heat flux of the GA-std case in D. . .	101

- 4.1 Demonstration of the calculation of  $k_{\max}$  and  $\gamma_{\max}$  from a linear growth rate spectrum, here for the  $a/L_n = 3.0$  case in H (section 4.6). Starting from  $\gamma_{k_y}$  (solid black), one divides this spectrum by  $k_y$  to obtain  $\gamma_{k_y}/k_y$  (dotted green). The maximum of this curve is then found via quadratic interpolation, with the  $k_y$  value at which this occurs being defined as  $k_{\max}$ . Returning then to  $\gamma_{k_y}$ , the value of this curve at  $k_{\max}$  is defined as  $\gamma_{\max}$ . These quantities are then used in the construction of  $\gamma_{\text{model}}$  (blue circles), given by equation 4.12. . . . . 109
- 4.2 Turbulent ion energy fluxes against density gradient scale length for (a) NL CGYRO, (b) TGLF-SAT1, (c) TGLF-SAT2 and (d) QuaLiKiz in H, D and T, from a GA-std case baseline. Note that the isotope scaling reversal present in NL CGYRO (shaded region) is not recreated in any of the quasilinear models. Subfigure (a) reproduced from [64], with the permission of AIP Publishing. . . . . 115
- 4.3 Turbulent ion energy flux against isotope mass for the GA-std case (a) and the  $a/L_n = 3.0$  case (b), for NL CGYRO (black), CGYRO-SAT1 (red), TGLF-SAT1 (blue), QuaLiKiz-qlk (yellow) and TGLF-SAT2 (green). Note the positive isotope scaling observed in both cases for the three fast quasilinear models, and the difference between the CGYRO-SAT1 and the TGLF-SAT1 results, originating solely from a difference in linear solver. . . . . 119
- 4.4 NL CGYRO ion flux components against wavenumber normalised to the main ion gyroradius  $k_y \rho_i$  for the  $a/L_n = 3.0$  case in H, D and T (a). In (b), the result of applying the  $\alpha_A$  metric at each  $k_y \rho_i$  is shown for the four models, as well as the value that would be seen if the fluxes followed gyroBohm scaling ( $\alpha_A = 0.5$ ). (c) Further isotope scaling results for NL CGYRO data from other cases in the database. The error bars shown for all models in (b) and (c) are the uncertainties in the fitted parameter  $\alpha_A$ . . . . . 120

4.5	CGYRO ion energy QLA function against $k_y\rho_i$ for the $a/L_n = 3.0$ case in H, D and T (a). In (b), the result of applying the $\alpha_A$ metric at each $k_y$ is shown, including the reference value one would expect if gyroBohm scaling were followed ( $\alpha_A = 0.0$ ). This value is expected as the scaling of the linear and nonlinear weights is the same, $W_{\text{GB}}^L, W_{\text{GB}}^{\text{NL}} \propto e^2 n_e c_s / T \propto A^{-0.5}$ , and thus their ratio cancels any predicted scaling dependence. The results of the quasilinear models are not shown as their QLA functions are taken to be constant, and thus exhibit no scaling ( $\alpha_A = 0.0$ ). (c) $\alpha_A$ results of the QLA functions for additional cases in the database. The error bars shown for (b) and (c) are the uncertainties in the fitted parameter $\alpha_A$ . . . . .	121
4.6	NL CGYRO squared potential magnitudes against $k_y\rho_i$ for the $a/L_n = 3.0$ case in H, D and T (a). In (b), the result of applying the $\alpha_A$ metric at each $k_y$ is shown, including the reference value one would expect if gyroBohm scaling were followed ( $\alpha_A = 1.0$ ), originating from $\delta\phi_{\text{GB}}^2 \propto \rho_*^2 (T/e)^2 \propto A$ . (c) Isotope scaling of the saturated potentials from NL CGYRO for other cases in the database. The error bars shown for all models in (b) and (c) are the uncertainties in the fitted parameter $\alpha_A$ . . . . .	122
5.1	Example of a $k_x$ -slice of the 2D potential magnitude spectrum, exhibiting the common features of an even, singularly peaked function with vanishing limits. Taken from the GA-std case in D, for $k_y\rho_{\text{unit}} = 10/39 = 0.256$ . . . . .	124
5.2	Measured values of $\sqrt{\sigma_{k_y}^2 C_1}$ against $k_y\rho_{\text{unit}}$ for the radially-extended ( $N_x = 896$ ) GA-standard case simulation in D, with $\sqrt{\sigma_{k_y}^2 C_1} = \sqrt{2}$ marked in red, corresponding to the squared Lorentzian assumption. . . . .	126
5.3	(a) An example of equation 5.7 (red) applied to NL CGYRO data (blue) for $k_y\rho_{\text{unit}} = 5/39 = 0.128$ of the radially-extended GA-std case, plotted against $k_x/\sigma_{k_y}$ . The squared Lorentzian model with the spectral width taken from the raw data is shown in green. (b) The same spectrum as (a) shown on a logarithmic $y$ -axis, with the predicted scaling regime limits marked. . . . .	127
5.4	(a) The model of equation 5.11 (black) applied to the NL CGYRO grid-independent potentials of the GA-standard case in D (red), allowing the two parameters to be fitted to the data. (b) Further examples of equation 5.11, for the $a/L_n = 3.0, \nu_{ee}(a/c_s) = 0.01$ case in H, D and T. . . . .	128

5.5	Measured values of the $c_1$ parameter across all nonlinear simulations in the database. The error bars shown are the uncertainties in the fitted parameter $c_1$ .	128
5.6	Example of $\sigma_{k_y}\rho_{\text{unit}}$ data obtained for the GA-standard case in D (blue). In red is the model for the widths given by equation 5.12, with the parameters $a, b$ and $c$ fitted to the data. Also marked are $k_{\text{min}}$ and $k_P$ , the positions of the minimum and piecewise connection point respectively.	130
5.7	Fitting exercise of the two quantities modelled as proportional to $k_{\text{max}}$ across the NL CGYRO database: $k_{\text{min}}$ (black circles), and $c/b$ (red squares), used in defining the model for the spectral shape $\sigma_{k_y}/\sigma_{k_y=k_0}$ (equation 5.15).	132
5.8	Potential magnitudes evaluated at $0.6k_{\text{min}}$ against a combination of linear quantities ((a) and (b)). Subfigure (c) displays the marker type used for each case, with the labels corresponding to those shown in table 4.1. Data points of the same colour connected by a line indicate simulations of the same equilibrium but of different isotopes. In (a), the cases with a dominant ITG instability scale approximately with $B_{\text{unit}}^2\gamma_{\text{max}}^2/k_{\text{max}}^5$ , however for those that are TEM-dominated (cases f, x, y and z, shown in (b)), a scaling of $B_{\text{unit}}^2\gamma_{\text{max}}^2\rho_{\text{unit}}/k_{\text{max}}^4$ is found to be in much better agreement, indicating a difference in saturation level between the two mode types.	133
5.9	Ratios of the linear electron energy weight magnitudes to the linear ion energy weight magnitudes $ W_{e,k_y}^L/W_{i,k_y}^L $ against $k_y/k_{\text{max}}$ for the database. Cases for which the dominant linear instability is the TEM are in blue, and those with ITG are shown in red. Note the disparate grouping between the mode types.	134
5.10	Linear energy weight magnitude ratios $ W_{e,k_y}^L/\sum_i W_{i,k_y}^L $ for different relative concentrations of D and T against $k_y/k_{\text{max}}$ for the GA-std case (a) and the $a/L_n = 3.0$ case (b). Note the essential invariance across the spectrum for the different concentrations.	136
5.11	QLA functions for all cases in the database, showing those for (a) ion energy flux, (b) electron energy flux and (c) particle flux.	139

5.12	Scatter plots of NL CGYRO against CGYRO-SAT3 results for (a) ion energy flux, (b) electron energy flux and (c) particle flux. The legend for these figures is shown in figure 5.8c. TEM-dominated cases are marked by circles (labelled f, x, y and z in figure 5.8c), with the remainder being ITG-dominated. Data points connected by a line denote simulations of the same equilibrium but different of isotopes. The line to denote perfect agreement between the models is shown in black, with the respective relative errors also shown on each plot, defined by equations 5.50 and 5.51. . . . .	143
5.13	Scatter plots equivalent to those of figure 5.12, however now comparing NL CGYRO against CGYRO-SAT1* results for (a) ion energy flux, (b) electron energy flux and (c) particle flux. . . . .	143
5.14	Turbulent ion energy fluxes (solid lines) and electron energy fluxes (dashed lines) against isotope mass for NL CGYRO (black), CGYRO-SAT3 (red) and CGYRO-SAT1* (blue), for the ITG-dominated GA-standard case (a) and the TEM-dominated $a/L_n = 3.0$ case (b). Note the recreation of the anti-gyroBohm scaling seen in the $a/L_n = 3.0$ case for CGYRO-SAT3. . . . .	144
5.15	(a) Ion energy fluxes for NL CGYRO (solid lines) and CGYRO-SAT3 (dashed lines) against density gradient scale length $a/L_n$ . (b) Ion energy fluxes for NL CGYRO and CGYRO-SAT3 against collisionality, keeping $a/L_n = 3.0$ fixed. (c) Ion and electron energy fluxes for NL CGYRO and CGYRO-SAT3 against matched temperature gradient scale length, $a/L_{T_i} = a/L_{T_e}$ . . . . .	145
5.16	(a) Isotope scaling metric $\alpha_A$ of the flux components for the $a/L_n = 3.0$ case, for NL CGYRO (black), CGYRO-SAT3 (purple), CGYRO-SAT1 (red), TGLF-SAT1 (blue) and QuaLiKiz-qlk (yellow). (b) The isotope scaling of the saturated potentials in the GA-std case for the same models. The error bars shown for all models are the uncertainties in the fitted parameter $\alpha_A$ . . . . .	145
6.1	Experimental time-traces of the L-mode discharges, with the dashed boxes indicating the window analysed in this chapter. From top to bottom, shown in (a) are the NBI powers and line-averaged densities, and in (b) are the gas puffing rates and total stored energy. Note the differing levels of NBI and puffing required to achieve the similarity in the stored energies and densities between the pulses. . . . .	150

6.2	Confinement time against line-averaged density for cases in a LOC-SOC scan for H, D and T. The four sets are grouped by connected circles, with sets 1, 2, 3 and 4 appearing in order of increasing density. Figure adapted from [145]. .	152
6.3	Values of the Ohmic power measured in EFIT, EFTP and TRANSP, with the corresponding JETTO interpretive simulations with $Z_{\text{eff}} = 1.05$ . Here the cases refer to those in increasing order of table 6.1, such that cases 1, 2 and 3 refer to the H, D and T shots of the low-density set. . . . .	154
6.4	Confinement time $\tau_E$ obtained for JETTO simulations using different quasilinear transport models for the set of L-mode pulses with (a) all effects included and (b) with no rotation. . . . .	157
6.5	Predicted electron ((a)-(c)) and ion ((d)-(f)) temperature profiles against major radius for the 3 L-mode simulations using various transport models. Note $T_i = T_e$ was assumed for the interpretive profiles. . . . .	159
6.6	Confinement time $\tau_E$ obtained for JETTO simulations using different quasilinear transport models, for a) the low-density set of Ohmic pulses and b) the mid-density set. . . . .	160
6.7	Confinement time $\tau_E$ obtained for JETTO simulations using different quasilinear transport models, for a) the first and b) the second high-density Ohmic sets. . . . .	161
6.8	Electron ((a)-(c)) and ion ((d)-(f)) temperature profiles against major radius for predictive JETTO simulations of the mid-density Ohmic set, with the experimental profiles shown in black. . . . .	162
6.9	Confinement time $\tau_E$ obtained for predictive JETTO simulations against line-averaged density $\bar{n}_e$ for the four turbulence models relative to the interpretive results. . . . .	163
E.1	(a) Normalised ion heat flux components against $k_y \rho_{\text{unit}}$ for the GA-std case in D, varying the resolution parameter $S$ . (b) The total ion heat fluxes for the cases in (a), plotted against $S$ . . . . .	219
E.2	(a) Normalised ion heat flux for the GA-std case in D against $k_y \rho_{\text{unit}}$ for different radial wavenumber resolutions. Note that the flux components are essentially unchanged between the simulations, indicating no issue of radial convergence. (b) A similar study for the $a/L_n = 3.0$ case in H, showing no appreciable change in the flux components across a ninefold change in radial resolution. .	220

# Acknowledgements

The completion of this thesis would not have been possible without the many people who have helped and supported me over the last four years. First and foremost I would like to thank my supervisors, Dr. Francis Casson and Dr. David Dickinson. Their combined wealth of knowledge and experience has been a continual source of inspiration throughout this research, and they have pushed me in becoming a better scientist. I am extremely grateful for their encouragement and insights in exploring this rich and wonderful subject.

During the course of this PhD I have been fortunate enough to benefit from several immensely rewarding collaborations, for which I have the pleasure of thanking Colin Roach, Gary Staebler, Emily Belli and Bhavin Patel. I am also grateful to Jonathan Citrin, Yann Camenen, Michael Fitzgerald, Costanza Maggi, Ephrem Delabie and Pablo Rodriguez-Fernandez for their useful insights and discussions, as well Bob Davies, Emma Hume and all my friends and colleagues on the Fusion CDT for their camaraderie both professionally and personally.

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. This work was supported by the Engineering and Physical Sciences Research Council [EP/L01663X/1, EP/R034737/1, EP/T012250/1, EP/W006839/1]. We appreciatively acknowledge access to the CINECA High-Performance Computer MARCONI, and to the JFRS-1 resource through QST.

Lastly I would like to thank my family, as well as James, Alexander and Robert, for providing me with that which is invaluable. I am forever indebted to you all.

# Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

A subset of the simulations performed in this work, outlined in Chapter 4, contributed to:

- G.M. Staebler, E.A. Belli, J. Candy, J.E. Kinsey, H. Dudding, and B. Patel, “Verification of a quasi-linear model for gyrokinetic turbulent transport”, *Nuclear Fusion* 61, 116007 (2021).

Aspects of Chapters 4 and 5 have been published in:

- H.G. Dudding, F.J. Casson, D. Dickinson, B.S. Patel, C.M. Roach, E.A. Belli and G.M. Staebler, “A new quasilinear saturation rule for tokamak turbulence with application to the isotope scaling of transport”, *Nuclear Fusion* 62, 096005 (2022).

# Chapter 1

## Introduction

### 1.1 Fusion energy

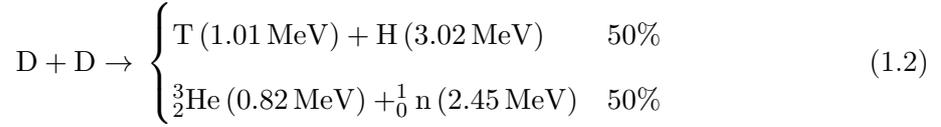
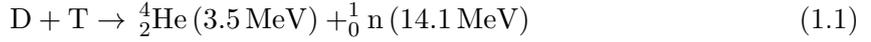
The need for humanity to secure reliable sources of energy for the future can hardly be overstated. The issue of continuing to meet rising global energy demands is compounded by the complications of both the depletion of the finite resources by which this energy is currently predominantly generated, and the negative environmental effects that are associated with their usage. A solution therefore requires a move towards renewable and green energy sources, such as wind and solar energy, however another candidate for the future energy mix is that of fusion energy.

Fusion energy relies on the principle of nuclear fusion, whereby light nuclei combine, with the resulting change in binding energy being released in the form of the energy of the reaction products. This is achieved by heating a gas to extremely high temperatures, such that its atoms become ionised and form a *plasma*, consisting of negatively charged electrons and positively charged ions. At such temperatures, characteristic of those in the core of stars, the average kinetic energy of the ions is such that a subset can overcome the electrostatic repulsion between them and come into sufficiently close proximity for the strong nuclear force to act. Fusion is therefore most easily achieved for isotopes of hydrogen, due to the relative weakness of the repulsion resulting from their atomic number  $Z = 1$ . The most important ions for fusion energy considerations are therefore standard hydrogen<sup>1</sup>, H, a single proton, deuterium D, consisting of a proton and a neutron, and tritium T, made up of a proton and two neutrons. Both H and D are stable isotopes, whereas T is weakly radioactive, with a half-life of 12.3 years [1]. The most important nuclear reactions for fusion energy considerations

---

<sup>1</sup>Sometimes referred to as 'protium'.

are [2]



where the partition of energy released between the fusion products is determined by conservation of momentum, and the percentages indicate the probabilities of each reaction branch occurring. The D-T fusion reaction represents the most advantageous, with the largest reaction rate (parameterised by the velocity-averaged reaction cross-section,  $\langle\sigma v\rangle$ ) occurring at a relatively low temperature<sup>2</sup> as seen in figure 1.1, and releasing a relatively large amount of energy per reaction at 17.6 MeV.

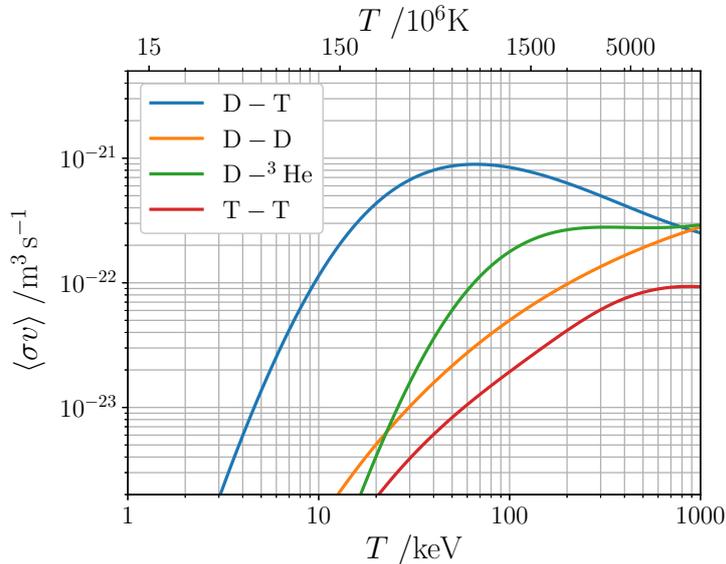


Figure 1.1: Relative velocity-averaged cross sections  $\langle\sigma v\rangle$  as a function of temperature for a selection of low  $Z$  fusion reactions. Figure adapted from [3], with the data obtained from [4].

Deuterium is a readily available resource accounting for approximately 0.0156% of the hydrogen atoms found in seawater [5]. The radioactivity of tritium renders it rare in nature, however it can be produced through neutron activation with lithium<sup>3</sup>. Estimates indicate that levels of naturally occurring lithium would allow for energy production on the order of thousands of years [6], and so the reactants of D-T fusion are in plentiful supply. Unlike the combustion of fossil fuels, the products from fusion reactions are not greenhouse gases, which

<sup>2</sup>Typically in plasma physics the temperature is expressed in dimensions of energy, with units of *electronvolts*, defined such that  $T[\text{eV}]e = T[\text{K}]k_B$ . Hence  $1 \text{ eV} \approx 1.16 \times 10^4 \text{ K}$ .

<sup>3</sup>The high energy neutron that results from D-T fusion can potentially be used to generate such reactions, which would allow fusion energy to sustain its own supply of tritium.

are associated with climate change. In comparison with nuclear fission, fusion energy features a far lower level of radioactivity.

On paper therefore, fusion appears to be the ideal energy source for the future, able to address the aforementioned issues of humanity's future energy needs. However despite being the subject of substantial research and development since the mid-twentieth century, no fusion experiment has yet occurred in which the energy produced as a consequence of the fusion reactions was greater than that required to generate them.

### 1.1.1 Fusion power balance

In order to analyse the conditions necessary to achieve net energy gain in a fusion reaction, zero-dimensional power balance will now be considered, similar to that first put forward by Lawson [7]. This is concerned with the rate of change of the stored energy of the plasma  $W = \int \sum_s \frac{3}{2} n_s T_s dV$ , where  $s$  labels the plasma species,  $n_s$  is the number density of species  $s$ ,  $T_s$  is the temperature and  $V$  is the volume of the plasma.

Fundamentally, the stored energy evolves according to the difference between the power imparted to the plasma to that lost,

$$\frac{dW}{dt} = P_{\text{in}} - P_{\text{out}}. \quad (1.3)$$

Fusion plasmas can receive energy from the charged products of the fusion reactions they generate, as they thermalise with the surrounding plasma. This heating can be estimated using the fusion reaction cross-section  $\sigma$ , for which it can be shown [8]

$$P_{\text{fus}} = \int n_1 n_2 \langle \sigma v \rangle E_{\text{ch}} dV \quad (1.4)$$

where  $n_1$  and  $n_2$  are the number densities of the fusing ions,  $E_{\text{ch}}$  is the energy of the charged products per fusion reaction and  $\langle \sigma v \rangle$  is the reaction rate averaged over velocity. Here we have assumed the charged products are confined for a sufficiently long time so as to completely thermalise, and that the neutral fusion products escape the plasma due to the lack of electromagnetic interaction. Assuming no impurities, then for ion number density  $n_i = n_1 + n_2$ , the product  $n_1 n_2$  is maximised for  $n_1 = n_2 = \frac{1}{2} n_i$ . For hydrogenic ions, the bulk neutrality of the plasma implies  $n_i = n_e$ , and hence one has  $n_1 n_2 = n_e^2 / 4$ .

One can also heat the plasma by external sources. These are included via an auxiliary heating term,  $P_{\text{aux}}$ .

To quantify the losses we introduce the *energy confinement time*  $\tau_E$ , defined such that

$$P_{\text{out}} = \frac{W}{\tau_E}. \quad (1.5)$$

Equation 1.3 now reads

$$\frac{dW}{dt} = P_{\text{aux}} + \int \frac{1}{4} n_e^2 \langle \sigma v \rangle E_{\text{ch}} dV - \frac{1}{\tau_E} W. \quad (1.6)$$

For energy gain to the plasma, one therefore requires

$$P_{\text{aux}} + \int \frac{1}{4} n_e^2 \langle \sigma v \rangle E_{\text{ch}} dV \geq \frac{1}{\tau_E} W \quad (1.7)$$

with the equality occurring in steady state. In the interest of energy generation, we would like to maximise our extractable energy out,  $P_{\text{ex}} = P_{\text{out}} + P_{\text{n}}$ , where  $P_{\text{n}}$  is the power imparted to the neutral fusion products, for minimal energy in. Assuming all extracted energy can be harnessed effectively, this efficiency can be quantified using the gain factor  $Q$ ,

$$\begin{aligned} Q &= \frac{P_{\text{ex}} - P_{\text{aux}}}{P_{\text{aux}}} \\ &= \frac{P_{\text{n}} + W/\tau_E - P_{\text{aux}}}{P_{\text{aux}}} \\ &= \frac{P_{\text{fus}}}{P_{\text{aux}}} \end{aligned} \quad (1.8)$$

where  $P_{\text{fus}} = P_{\text{ch}} + P_{\text{n}}$ . The case of  $Q = 1$  corresponds to the point of energy breakeven<sup>4</sup>.

## Ignition

Of particular interest is the condition for which the auxiliary heating can be switched off, and the fusion reactions be sustained solely by charged particle heating. This corresponds to  $Q \rightarrow \infty$  and is called *ignition*. As a simple estimation of the conditions necessary, we set  $P_{\text{aux}} = 0$  in equation 1.7 to find

$$\int \frac{1}{4} n_e^2 \langle \sigma v \rangle E_{\text{ch}} dV \geq \frac{1}{\tau_E} \int 3n_e T dV \quad (1.9)$$

which re-arranges to

$$n_e \tau_E \geq \frac{12T}{\langle \sigma v \rangle E_{\text{ch}}}. \quad (1.10)$$

For a Maxwellian distribution, the quantity  $\langle \sigma v \rangle / T^2$  contains a global maximum, as seen in figure 1.2, which for D-T fusion is  $\langle \sigma v \rangle / T^2 \approx 1.1 \times 10^{-24} \text{ m}^3 \text{ s}^{-1} \text{ keV}^{-2}$  applicable over the range  $T = 10 - 20 \text{ keV}$ . Using  $E_{\text{ch}} = 3.5 \text{ MeV}$  from equation 1.1, the ignition condition for D-T fusion may therefore be approximated as

$$n_e T \tau_E \geq 3 \times 10^{21} \text{ keV s m}^{-3} \quad (1.11)$$

---

<sup>4</sup>For reactor considerations we note that  $P_{\text{aux}}$  only counts the power delivered to the plasma, not the energy spent in the operation of the auxiliary heating methods. The gain factor that takes this effect into account is the ‘engineering gain’  $Q_{\text{eng}}$ , and is smaller than the value of  $Q$  in equation 1.8, which becomes smaller still when including losses from methods of subsequent energy generation.

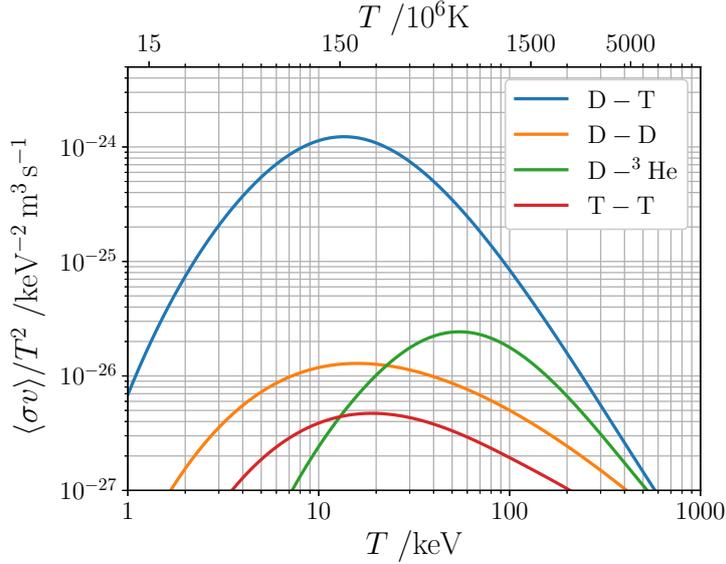


Figure 1.2: The data of figure 1.1 now divided by the square of the temperature. Note the appearance of a global maximum for each reaction.

where the left-hand side is known as the *fusion triple product*, and can be used as a figure of merit for a given fusion experiment. While this simple analysis has ignored a number of effects such as radiation and impurities, it illustrates the central concept of fusion power: one must confine a hot and dense plasma for a sufficiently long time, so as to allow reactions yielding net energy to be sustained. For fusion-relevant temperatures, and assuming a density of  $n \sim \mathcal{O}(10^{20} \text{ m}^{-3})$ , this corresponds to a confinement time of  $\tau_E \sim \mathcal{O}(1 \text{ s})$ . In order to describe how one would confine a plasma in such a way, we first consider some of their physical properties.

## 1.2 Physics of plasmas

### 1.2.1 Plasma models

A fusion plasma can be modelled as an ideal gas, consisting of electrons and one or more ion species, which at fusion-relevant temperatures can be taken to be fully ionised. Each species is characterised by its mass  $m_s$  and electric charge  $q_s = Z_s e$ , as well as its number density  $n_s$  and temperature  $T_s$ , which are defined below. A typical speed may then be defined using the thermal velocity,  $v_{\text{th},s} = \sqrt{T_s/m_s}$ .

## Single particle description

Considering each particle in the plasma individually, the dynamics are fundamentally determined by the set of coupled equations

$$\frac{d\mathbf{r}_j}{dt} = \mathbf{v}_j \quad (1.12)$$

$$\frac{d\mathbf{v}_j}{dt} = \frac{q_j}{m_j} (\mathbf{E} + \mathbf{v}_j \times \mathbf{B}) \quad (1.13)$$

$$\nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0} \rho \quad (1.14)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (1.15)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (1.16)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad (1.17)$$

where equation 1.13 is Newton's second law with the Lorentz force, the label  $j$  applies to every individual particle in the plasma, and the charge density and current density are calculated from the ensemble of particles. Treating each particle in the plasma as a system of coupled equations is hypothetically the most complete, however is impractical in the numbers of particles typically considered<sup>5</sup>.

## Statistical description

Instead, one can consider a statistical description. This describes a species in a plasma in terms of its *distribution function*  $f_s(t, \mathbf{r}, \mathbf{v})$ , which evolves in phase space via  $df_s/dt = C_s$ , where  $C_s$  is the collision operator<sup>6</sup>. Using the chain rule, one obtains the kinetic equation [9]

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \nabla f_s + \frac{q_s}{m_s} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_s}{\partial \mathbf{v}} = C_s. \quad (1.18)$$

This equation is typically referred to as the 'Vlasov equation' in the absence of collisions. The distribution function  $f_s$  describes the behaviour of the ensemble of a given species in 6D phase space and time, with velocity moments of  $f_s$  corresponding to physical observables. Most important for our considerations are the first three velocity moments, defining the number density, average velocity, and temperature

$$n_s(\mathbf{r}, t) = \int f_s d^3\mathbf{v} \quad (1.19)$$

---

<sup>5</sup>For confinement devices known as 'tokamaks', discussed in section 1.4, a typical plasma density of  $n \sim \mathcal{O}(10^{20} \text{ m}^{-3})$  confined to a volume  $V \sim \mathcal{O}(10^2 \text{ m}^3)$  provides an estimate of  $\sim 10^{22}$  particles.

<sup>6</sup>The exact form of the collision operator will not be considered in this work, instead being treated schematically.

$$n_s \mathbf{u}_s(\mathbf{r}, t) = \int \mathbf{v} f_s d^3\mathbf{v} \quad (1.20)$$

$$\frac{3}{2} n_s T_s(\mathbf{r}, t) = \frac{1}{2} m_s \int (\mathbf{v} - \mathbf{u}_s)^2 f_s d^3\mathbf{v}. \quad (1.21)$$

The charge density and current density present in Maxwell's equations can thus be found from

$$\rho(\mathbf{r}, t) = \sum_s q_s \int f_s d^3\mathbf{v} \quad (1.22)$$

$$\mathbf{J}(\mathbf{r}, t) = \sum_s q_s \int \mathbf{v} f_s d^3\mathbf{v}. \quad (1.23)$$

The kinetic system of plasma description is made up of equations 1.14-1.17 and one kinetic equation (1.18) for each particle species present. Even having replaced the description of each individual particle with a single equation describing all particles of a given species, the kinetic system can still be prohibitively costly with all but supercomputers, owing to the high dimensionality of the problem, as well as the breadth of spatial, temporal and velocity scales present in general plasma dynamics<sup>7</sup>.

## Fluid system

By taking velocity moments of the kinetic equation, one can derive a system of equations which describes the evolution of the previously-defined observables. This reduces the dimensionality of the plasma description from 7 to 4, alleviating much computational expense from the problem, however comes at the cost of increasing the number of equations one is required to evolve. These are known as the fluid equations, and their derivation is demonstrated in appendix A.1. The zeroth moment of equation 1.18 is

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{u}_s) = 0 \quad (1.24)$$

which is the continuity equation for the particles of species  $s$ <sup>8</sup>. The first velocity moment, which describes the evolution of momentum, is

$$m_s n_s \left[ \frac{\partial \mathbf{u}_s}{\partial t} + (\mathbf{u}_s \cdot \nabla) \mathbf{u}_s \right] + \nabla P_s + \nabla \cdot \underline{\underline{\pi}}_s - q_s n_s (\mathbf{E} + (\mathbf{u}_s \times \mathbf{B})) = \int m_s \mathbf{v} C_s d^3\mathbf{v} \quad (1.25)$$

where  $P_s = n_s T_s$  is the scalar pressure and  $\underline{\underline{\pi}}_s$  is the anisotropic pressure tensor. One may continue this operation to obtain the second moment equation, describing the evolution of the pressure, and then the third, and so on. While this does reduce the dimensionality of the

<sup>7</sup>A discussion of the disparate scales relevant to the considerations of this work is presented in section 3.1.2.

<sup>8</sup>Note that multiplying this equation by the species charge  $q_s$  and summing over species produces the equation of charge conservation, as is encoded in Maxwell's equations.

problem, we are confronted with the issue of the fluid hierarchy. For every moment equation derived, the evolution of the  $n^{\text{th}}$  moment is described in terms of the  $n + 1^{\text{th}}$ . We therefore need to introduce an approximation to close the system [10, 11]. Such closures can vary in complexity but cannot be self-consistent, and so can limit the regime of applicability of a purely fluid description.

The decision regarding which plasma models to use, namely the motion of single particles, a fluid or kinetic description, or a combination of models is largely dictated by the problem at hand. For example, in the consideration of large scale plasma instabilities, the framework of magnetohydrodynamics (MHD) [6], which treats the plasma as a single conducting fluid, is commonly employed. In this work all three are used for different considerations. We now consider some plasma concepts relating to these models.

### 1.2.2 Adiabatic response

A common simplifying assumption that can be made to the dynamics of the electrons, given their disparate masses relative to the ions, is the *adiabatic response*<sup>9</sup>. In this case, the inertia term is assumed to be small relative to the other terms in the electron momentum equation (equation 1.25), such that parallel to the field, only the pressure fluctuations and the electric field fluctuations play a significant role.

Linearising the component of the momentum equation parallel to the magnetic field and assuming isothermal electrons, one finds

$$T_e \nabla_{\parallel} \delta n_e = e n_{e,0} \nabla_{\parallel} \delta \phi \quad (1.26)$$

where  $\nabla_{\parallel} = \hat{\mathbf{b}} \cdot \nabla$ , for unit vector in the direction of the magnetic field  $\hat{\mathbf{b}}$ , and  $n_e = n_{e,0} + \delta n_e$ .

This integrates to

$$\delta n_e = \frac{e n_{e,0}}{T_e} \delta \phi \quad (1.27)$$

where the constant of integration has been set by the condition  $\delta \phi = 0 \implies \delta n_e = 0$ . Equation 1.27 is the adiabatic response, in which the electrons respond instantaneously to a change in potential via a change in density, and is often a useful assumption to lowest order.

### 1.2.3 Quasineutrality

On a large scale plasmas are electrically neutral, due to the balance between the number of electrons and the positive charges of the ions,  $n_e = \sum_i Z_i n_i$ . On a small scale however, local

---

<sup>9</sup>Sometimes known as the ‘Boltzmann response’.

regions of net charge can occur due to the motion of particles. The exact meanings of ‘large’ and ‘small’ in this context are described by the concept of *quasineutrality*. By considering the insertion of an additional point charge into an otherwise neutral plasma, one can use Gauss’s law to estimate the electrostatic potential  $\phi$  around the particle, and hence the length scale over which this effect is relevant. For distance  $r$  centered on this point particle, one finds a potential of the form (appendix A.2)

$$\phi(r) = \frac{1}{4\pi\epsilon_0} \frac{e^{-r/\lambda_D}}{r} \quad (1.28)$$

where  $\lambda_D = \sqrt{\epsilon_0 T_e / n_e e^2}$  is the *Debye length*. This is simply the vacuum potential of a point charge multiplied by an exponential decay over distance. The potential is ‘screened’ by the plasma, such that it is negligible on length scales larger than the Debye length. For a given length scale  $l$ , the charge density therefore follows

$$\begin{cases} \rho \neq 0 & l \ll \lambda_D \\ \rho = 0 & l \gg \lambda_D \end{cases} \quad (1.29)$$

and thus a plasma can be considered electrically neutral on length scales  $l \gg \lambda_D$ .

#### 1.2.4 Particle motion in uniform electric and magnetic fields

Although plasmas can be considered neutral on a large scale, they are fundamentally made up of charged particles, each of which responds to electric and magnetic fields. Analysis of the case of constant fields is given here.

We first consider a Cartesian coordinate system in which the magnetic field is aligned with the  $z$ -axis,  $\mathbf{B} = B\hat{\mathbf{k}}$ . For the unit vector in the direction of the magnetic field  $\hat{\mathbf{b}} = \mathbf{B}/B$ , in this instance we simply have  $\hat{\mathbf{b}} = \hat{\mathbf{k}}$ . A general vector  $\mathbf{V}$  may then be split into components parallel and perpendicular to the field, such that  $\mathbf{V} = V_{\parallel}\hat{\mathbf{b}} + \mathbf{V}_{\perp}$ , defined via  $V_{\parallel} = \mathbf{V} \cdot \hat{\mathbf{b}}$  and  $\mathbf{V}_{\perp} = \mathbf{V} - V_{\parallel}\hat{\mathbf{b}}$ . Without loss of generality, we align the perpendicular component of the electric field with the  $y$ -axis,  $\mathbf{E} = E_{\parallel}\hat{\mathbf{b}} + E_{\perp}\hat{\mathbf{j}}$ . In the constant field case for a single particle of species  $s$ , the dynamical equations are

$$\frac{d\mathbf{v}}{dt} = \frac{q_s}{m_s} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (1.30)$$

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}. \quad (1.31)$$

The parallel component of equation 1.30 is

$$\frac{dv_{\parallel}}{dt} = \frac{q_s}{m_s} E_{\parallel} \quad (1.32)$$

which can be integrated to obtain

$$v_{\parallel}(t) = v_{\parallel}(0) + \frac{q_s}{m_s} E_{\parallel} t. \quad (1.33)$$

This result simply states that the parallel motion is only affected by the parallel electric field component. For the case of  $E_{\parallel} = 0$ , the particle moves at a constant speed parallel to the magnetic field.

Returning to equation 1.30, then moving the  $\mathbf{v} \times \mathbf{B}$  term to the left-hand side, the perpendicular component reads

$$\frac{d\mathbf{v}_{\perp}}{dt} - \left( \frac{q_s B}{m_s} \right) \mathbf{v}_{\perp} \times \hat{\mathbf{b}} = \frac{q_s}{m_s} \mathbf{E}_{\perp}. \quad (1.34)$$

Expressing the velocity in terms of a solution to the homogeneous equation  $\mathbf{u}$  and a particular integral  $\bar{\mathbf{v}}_{\perp}$  such that  $\mathbf{v}_{\perp} = \mathbf{u} + \bar{\mathbf{v}}_{\perp}$ , we consider first the homogeneous equation,

$$\frac{d\mathbf{u}}{dt} - \left( \frac{q_s B}{m_s} \right) \mathbf{u} \times \hat{\mathbf{b}} = \mathbf{0} \quad (1.35)$$

which upon differentiating with time yields

$$\frac{d^2\mathbf{u}}{dt^2} + \Omega_s^2 \mathbf{u} = \mathbf{0} \quad (1.36)$$

where  $\Omega_s = q_s B / m_s$ . This is an equation of simple harmonic motion for the velocity vector  $\mathbf{u}$ , with angular frequency  $\Omega_s$ , known as the *gyrofrequency*<sup>10</sup>. Taking the components of this vector equation, the general solutions are  $u_x = A_u \cos(\Omega_s t + \phi_u)$  and  $u_y = -A_u \sin(\Omega_s t + \phi_u)$ , where the amplitude  $A_u^2 = u_x^2 + u_y^2 = |\mathbf{u}|^2$  is a constant, and  $\tan(\phi_u) = -u_y(0) / u_x(0)$ . Without loss of generality we choose the positive root of  $A_u = |\mathbf{u}|$ , and  $\phi_u = 0$ . Therefore the velocity vector  $\mathbf{u}$  undergoes circular motion at the gyrofrequency,  $\Omega_s$ , given by

$$\mathbf{u} = |\mathbf{u}| \left( \cos(\Omega_s t) \hat{\mathbf{i}} - \sin(\Omega_s t) \hat{\mathbf{j}} \right). \quad (1.37)$$

We now consider the inhomogeneous equation. Because the right-hand side is a constant vector, assuming the solution to be constant gives

$$-\Omega_s \bar{\mathbf{v}}_{\perp} \times \hat{\mathbf{b}} = \frac{q_s}{m_s} \mathbf{E}_{\perp} \quad (1.38)$$

which, by taking the cross product with  $\hat{\mathbf{b}}$ , rearranges to

$$\bar{\mathbf{v}}_{\perp} = \frac{\mathbf{E}_{\perp} \times \hat{\mathbf{b}}}{B} \quad (1.39)$$

which is known as the ' $E \times B$ ' velocity,  $\mathbf{v}_E$ . This causes the particle to drift at a constant velocity in a direction perpendicular to both the electric field and the magnetic field.

---

<sup>10</sup>Also known as the cyclotron frequency.

Collecting these three parts of the motion together, the velocity of a charged particle in constant electric and magnetic fields is

$$\mathbf{v}(t) = \left[ v_{\parallel}(0) + \frac{q}{m} E_{\parallel} t \right] \hat{\mathbf{b}} + \frac{\mathbf{E}_{\perp} \times \hat{\mathbf{b}}}{B} + |\mathbf{u}| \left( \cos(\Omega_s t) \hat{\mathbf{i}} - \sin(\Omega_s t) \hat{\mathbf{j}} \right) \quad (1.40)$$

which, upon integrating to solve for the particle position, gives

$$\begin{aligned} \mathbf{r}(t) &= \mathbf{r}(0) + \int_0^t \mathbf{v}(t') dt' \\ &= \bar{\mathbf{r}}(0) + \left[ v_{\parallel}(0) t + \frac{1}{2} \frac{q_s}{m_s} E_{\parallel} t^2 \right] \hat{\mathbf{b}} + \frac{\mathbf{E}_{\perp} \times \hat{\mathbf{b}}}{B} t + \rho \left[ \sin(\Omega_s t) \hat{\mathbf{i}} + \cos(\Omega_s t) \hat{\mathbf{j}} \right] \end{aligned} \quad (1.41)$$

where  $\bar{\mathbf{r}}(0) = \mathbf{r}(0) - \frac{|\mathbf{u}|}{\Omega_s} \hat{\mathbf{j}}$  is related to the initial position of the particle, and  $\rho = |\mathbf{u}|/\Omega_s$  is the *gyroradius*. Equation 1.41 expresses the motion of a charged particle as a superposition of three parts: parallel streaming along the field lines, the perpendicular  $E \times B$  drift, and the *gyromotion* of the particle, consisting of circular motion at constant frequency  $\Omega_s$  and constant radius  $\rho$ . The angle of the particle along this orbit is the *gyrophase*,  $\varphi = \Omega_s t$ . Expressing the gyromotion term as a vector  $\boldsymbol{\rho}$ , we note its relation to the velocity of the gyromotion

$$\boldsymbol{\rho} = \frac{|\mathbf{u}|}{\Omega_s} \left[ \sin(\Omega_s t) \hat{\mathbf{i}} + \cos(\Omega_s t) \hat{\mathbf{j}} \right] = \frac{\hat{\mathbf{b}} \times \mathbf{u}}{\Omega_s}. \quad (1.42)$$

Note that the gyrofrequency  $\Omega_s$  is positive for ions and negative for electrons due to the factor of  $q_s$  in its definition. This causes them to rotate around magnetic field lines in opposite orientations, with ions orbiting in a clockwise direction and electrons in an anticlockwise direction.

For the case of no electric field, one has

$$\mathbf{r}(t) = \bar{\mathbf{r}}(0) + v_{\parallel}(0) t \hat{\mathbf{b}} + \rho \left[ \sin(\Omega_s t) \hat{\mathbf{i}} + \cos(\Omega_s t) \hat{\mathbf{j}} \right] \quad (1.43)$$

and thus we see that the motion of the particle traces out a helix, consisting of constant parallel streaming combined with gyromotion perpendicular to the field, as illustrated in figure 1.3. We note that in this simple case, this essentially confines the particle to a single magnetic field line, restricting the motion of the particle in the perpendicular directions.

The presence of an  $E \times B$  drift causes the particle to no longer be tied to a single magnetic field line, but to drift across them, as demonstrated in figure 1.4. Note that the  $E \times B$  velocity is *independent* of charge and mass, and thus ions and electrons drift in the same direction with the same speed.

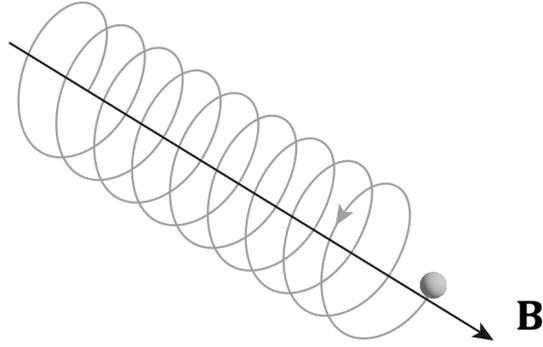


Figure 1.3: Particle gyromotion in the presence of a uniform magnetic field and no electric field. Note that the anti-clockwise orientation of the motion corresponds to an electron.

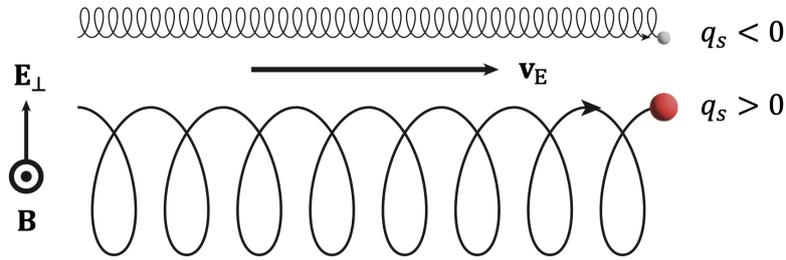


Figure 1.4:  $E \times B$  motion of an electron (grey) and ion (red) in the presence of constant orthogonal electric and magnetic fields. The dynamics parallel to the magnetic field have been suppressed for simplicity.

### The gyrocentre

For many dynamical considerations one is not so concerned as to the exact position of the particle with respect to its gyromotion, but the motion of the particle due to its parallel motion and drifts. For this purpose we define the instantaneous centre of the gyromotion as the *gyrocentre*  $\mathbf{R}(t) = \mathbf{r}(t) - \boldsymbol{\rho}(t)$ . The position of the gyrocentre in the constant field case is therefore

$$\mathbf{R}(t) = \mathbf{R}(0) + \left[ v_{\parallel}(0)t + \frac{1}{2} \frac{q_s}{m_s} E_{\parallel} t^2 \right] \hat{\mathbf{b}} + \frac{\mathbf{E}_{\perp} \times \hat{\mathbf{b}}}{B} t. \quad (1.44)$$

The relation of these vectors is shown in figure 1.5.

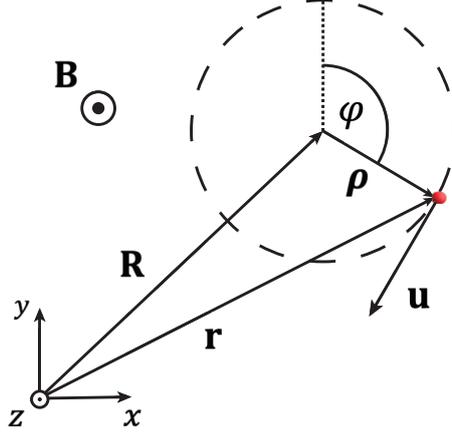


Figure 1.5: Schematic illustrating the relation between the particle position  $\mathbf{r}$ , the gyrocentre  $\mathbf{R}$ , the gyroradius  $\rho$  and the gyrophase  $\varphi$ .

### Conserved quantities

We have already noted that the magnitude of the velocity vector associated with gyromotion  $|\mathbf{u}|$  is conserved. The system contains no positional dependence, and so any function of the  $\mathbf{E}$  and  $\mathbf{B}$  fields is trivially conserved. We also note that by taking the dot product of the velocity with the Lorentz force equation, one has

$$\mathbf{v} \cdot \frac{d\mathbf{v}}{dt} = \frac{q_s}{m_s} \mathbf{E} \cdot \mathbf{v} \quad (1.45)$$

which, when using the electrostatic form of the electric field from Faraday's law,  $\mathbf{E} = -\nabla\phi$ , one can use  $d\phi/dt = \partial\phi/\partial t + \mathbf{v} \cdot \nabla\phi$  to write

$$\mathbf{v} \cdot \frac{d\mathbf{v}}{dt} = -\frac{q_s}{m_s} \mathbf{v} \cdot \nabla\phi \implies \frac{d}{dt} \left( \frac{1}{2} m_s |\mathbf{v}|^2 + q_s \phi \right) = 0 \quad (1.46)$$

where  $\partial\phi/\partial t = 0$  in the constant field case. We hence find that the particle energy  $\varepsilon = \frac{1}{2} m_s |\mathbf{v}|^2 + q_s \phi$  is conserved in the constant field case, as expected.

## 1.3 Plasma confinement

The foregoing concepts on the physics of plasmas can now be used to discuss possible methods by which one is able to confine them, in pursuit of a sufficiently large value of  $\tau_E$  for the purposes of satisfying equation 1.11. Due to the high temperatures required to facilitate fusion reactions, any physical container would be instantly damaged upon contact with such an extreme environment, and thus cannot be used. By exploiting the electrically-charged nature of plasmas, one can instead consider a confinement method using a magnetic field,

based on the concept of ‘tying’ the particles to field lines via their gyromotion. Confinement methods based on this principle come under the heading of magnetic confinement fusion.

Initially considering a plasma in vacuum, one can envision that by passing a strong external magnetic field through the plasma, each particle will gyrate around the magnetic field lines, restricting their motion and confining the plasma in the two dimensions perpendicular to the field. This concept must however address two issues. Firstly, while this method provides a basis of confinement for the two dimensions perpendicular to the magnetic field, this says nothing of confinement parallel to the field, for which the particles will stream along the field lines by their thermal motion. The second issue is that magnetic fields are rarely uniform in practice, and typically contain some degree of spatial and temporal variation. This variation in the field generates additional particle drifts, associated with the curvature and inhomogeneity in the field strength. Assuming that these variations are on a scale much larger than the ion gyroradius,  $B/|\nabla B| \gg \rho_i$ , these can be shown to be<sup>11</sup>

$$\mathbf{v}_{D,s} = \frac{v_{\parallel}^2}{\Omega_s} \hat{\mathbf{b}} \times \left[ \left( \hat{\mathbf{b}} \cdot \nabla \right) \hat{\mathbf{b}} \right] + \frac{|\mathbf{u}|^2}{2\Omega_s} \hat{\mathbf{b}} \times \nabla B \quad (1.47)$$

which are the ‘curvature drift’ and the ‘grad-B drift’ respectively. Note that both of these drifts are perpendicular to the magnetic field, and thus have an effect similar to the  $E \times B$  motion seen previously, however now with a dependence on charge and mass through the factor of  $\Omega_s$ .

Returning to the issue of the parallel confinement, one can consider the configuration of bending the magnetic field lines such that they return back on themselves, forming circular magnetic field lines around an axis, as illustrated in figure 1.6. One could, for example, consider passing a current  $I$  through a vertical wire, generating a concentric magnetic field. That way, one would hope that the particles would continue to stream round in a circle indefinitely, while remaining confined in the dimensions perpendicular to the field, and thus form a basis for confinement. However, such a configuration has introduced curvature to the field, which also incurs inhomogeneity in the field magnitude, resulting in drifts given by equation 1.47. Considering a cylindrical coordinate system  $\{R, \Phi, Z\}$ , this would generate a field described by  $\mathbf{B} = B(R) \hat{\mathbf{e}}_{\Phi}$ , with  $dB/dR < 0$ .

Using this form of the field to calculate the direction of the drifts in equation 1.47, these are found to be in the  $+\hat{\mathbf{k}}$  direction for ions and in the  $-\hat{\mathbf{k}}$  direction for electrons, owing to the factor of the gyrofrequency  $\Omega_s$  present in both drifts. The two particle types would therefore drift vertically in *opposite* directions as a result of this configuration, generating a separation

---

<sup>11</sup>These drifts are derived in Section 3.2.

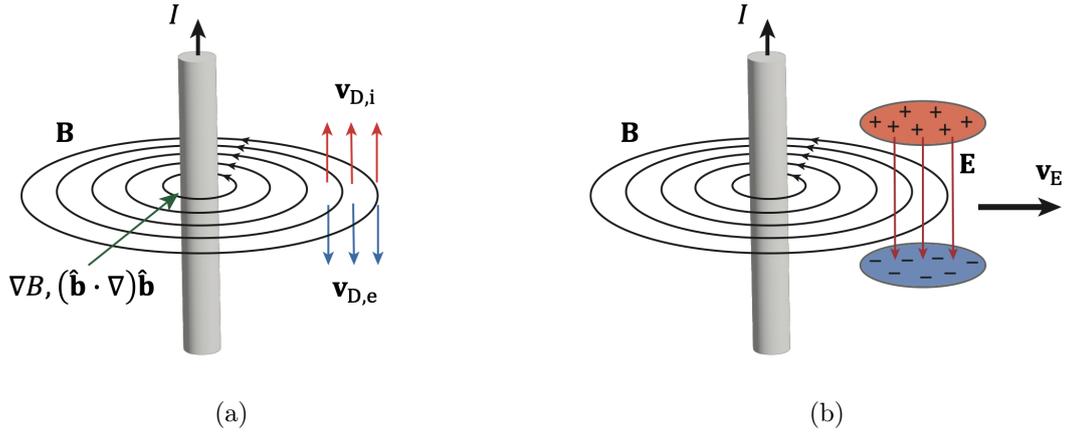


Figure 1.6: Exploring the case of attempting plasma confinement using a concentric circular magnetic field. The drifts resulting from the field configuration would cause a separation of space charge between the ions and electrons, generating an  $E \times B$  drift which would cause confinement to be lost.

of charge and thus an electric field. The resulting  $E \times B$  drift would then be directed radially outwards, causing confinement to be lost unacceptably quickly [12].

## 1.4 Tokamaks

To remedy this loss of confinement, a mechanism must be included such that any charge separation resulting from the different drifts can be effectively shorted. This can be achieved by the inclusion of an additional magnetic field, such as to allow particles to traverse along the field between the vertical extremities. Considering the field shown in figure 1.6 to be the long way around a torus (the ‘toroidal’ direction), the inclusion of a magnetic field component in the ‘poloidal’ direction (the short way around a torus) resolves this issue, allowing a route by which particles can stream parallel to the field to short out potential charge separation. The configuration of a toroidal field and poloidal magnetic field can be broadly split into two types of devices, depending on the variation of the machine in the toroidal angle direction. Those that are *axisymmetric* are tokamaks, and those that include toroidal variation are called stellarators<sup>12</sup>.

<sup>12</sup>Stellarators are not considered in this work.

## Magnetic field configuration

The way in which tokamaks generate their magnetic configuration is demonstrated in figure 1.7. The toroidal field of a tokamak is generated using a set of external toroidal field coils. A current is then passed through a central solenoid, inducing via transformer action a toroidal current in the plasma<sup>13</sup>. This current generates the poloidal magnetic field through Ampère's law. Additional external coils can be included to alter the magnetic field configuration. The uniformity with which the toroidal field coils are arranged around the chamber causes the tokamak to be axisymmetric, exhibiting minute variation in the toroidal direction.

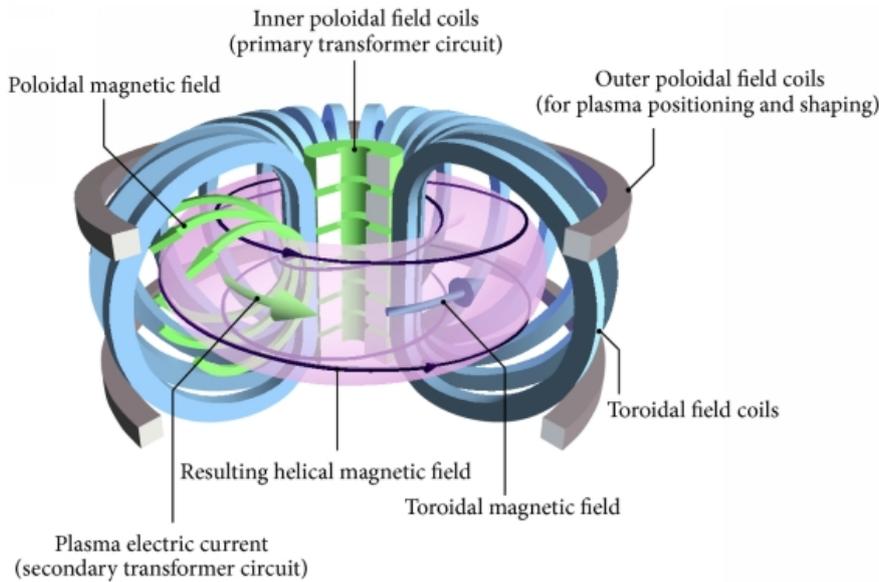


Figure 1.7: Schematic of the magnetic field configuration in a tokamak [13].

## Tokamak geometry

The principal aspects of the tokamak configuration will now be considered<sup>14</sup>. The combination of the poloidal and toroidal fields with the condition of axisymmetry causes the toroidal geometry to consist of concentric, nested surfaces of constant plasma pressure  $P$ , shown in figure 1.8. These are called *flux surfaces*, and their labelling forms a natural radial coordinate for the tokamak,  $\rho$ . Also shown in figure 1.8 is the major radius  $R_0$  which measures from the central axis of the solenoid to the magnetic axis, the minor radius  $a$  which measures from the magnetic axis to the edge of the plasma, and the poloidal and toroidal angles  $\vartheta$  and  $\zeta$ . The

<sup>13</sup>Non-inductive current drive approaches may also be considered, in part due to the inherent challenge of obtaining a steady-state plasma discharge through the use of inductive current drive.

<sup>14</sup>These results are derived in Section 3.4.

ratio of the major radius to the minor radius defines the *aspect ratio* of the tokamak,  $R_0/a$ . Standard ‘doughnut’ shaped configurations, such as that shown in figure 1.8, typically have aspect ratios  $\sim 3$  and are called ‘large aspect ratio’ tokamaks<sup>15</sup>.

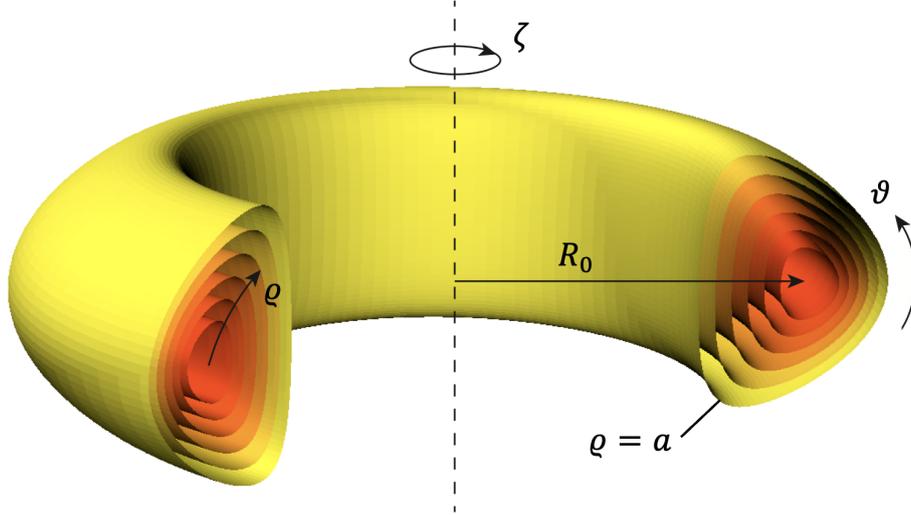


Figure 1.8: Schematic of flux surfaces present in a tokamak, with coordinate labels for the radial, poloidal and toroidal directions  $\{\varrho, \vartheta, \zeta\}$ . Figure adapted from [14].

Quantities that are constant on a given flux surface are called *flux functions*. One such flux function is the equilibrium plasma pressure  $P = P(\varrho)$ , which is largest towards the centre of the poloidal cross-section (the ‘core’) and decreases radially towards the ‘edge’. In the limit of small plasma rotation, on each flux surface the pressure gradient is balanced by the equilibrium  $\mathbf{J} \times \mathbf{B}$  force,

$$\mathbf{J} \times \mathbf{B} = \nabla P. \quad (1.48)$$

This condition of equilibrium force balance implies  $\mathbf{B} \cdot \nabla P = \mathbf{J} \cdot \nabla P = 0$ , and thus both the equilibrium magnetic field and the equilibrium current density lie entirely on these flux surfaces.

The combination of the toroidal and poloidal magnetic field results in a helical field configuration on the flux surfaces. The average pitch angle of the magnetic field is related to the ‘safety factor’,  $q(\varrho)$ , named due to its influence on the macroscopic stability of plasmas [6]. This approximately measures the ratio of the number of turns in the toroidal direction per

<sup>15</sup>This is in contrast to small aspect ratio tokamaks, or ‘spherical tokamaks’ (STs), which are characterised by an aspect ratio closer to 1, and resemble more of a cored apple than a doughnut. In this work we will focus on large aspect ratio tokamaks.

turn in the poloidal direction for a given flux surface,

$$q(\varrho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\mathbf{B} \cdot \nabla \zeta}{\mathbf{B} \cdot \nabla \vartheta} d\vartheta. \quad (1.49)$$

### Experimental tokamaks

While the analysis in this chapter has expounded the concept of magnetic confinement for tokamaks in principle, how effective is the confinement achieved by this configuration in practice?

Since their conception in the late 1950's, tokamaks have been the forerunner for a fusion reactor design, and are the most advanced form of fusion device. They have seen steady levels of R&D, with tokamaks of various shapes and sizes existing all over the world. In all this time however, no tokamak has yet achieved energy breakeven,  $Q = 1$ . The current record for the fusion energy gain factor was set by the tokamak JET's first D-T campaign in 1996 of  $Q = 0.62$  [15]. The relative scarcity of tritium, combined with the difficulties in its utilisation due to its radioactivity, has meant that D-T reactions have been rare in the history of tokamak experiments, with the overwhelming majority operating using D-D fusion. While this reaction is not the optimal one for fusion energy, the confinement times  $\tau_E$  measured for all experiments are drastically lower than initial, more optimistic estimates of early toroidal fusion designs [16].

The poor confinement properties of devices are due to the presence of *turbulence* in tokamaks, which degrades the confinement of the plasma by enhancing the radial transport of particles and energy down the pressure gradient, away from the tokamak core and out past the edge. The issue of plasma turbulence remains one of the outstanding problems in fusion physics, and its study forms a large part of modern fusion science. The effects of the resulting radial transport is the remit of confinement physics, and is considered in the next chapter.

Notwithstanding the issue of turbulence and its associated degradation of confinement, the fusion community currently awaits the arrival of the next-generation tokamak experiment ITER [17], currently under construction in the south of France. This machine represents a collaboration between many of the world's nations to demonstrate the feasibility of fusion power, and will be the largest tokamak to date, with the increase in size of the device expected to significantly improve fusion performance.

Purpose built for D-T, the goals of ITER [18] include producing brief plasma pulses of  $Q = 10$  and steady-state pulses of  $Q \geq 5$ . While ITER will not produce any of its own electricity, it will provide a testing facility to experiment with the required technologies for future devices.

The success of ITER therefore represents a proof of concept of the feasibility of fusion power. A sizeable portion of the efforts of the international fusion community has been to prepare for the physics and engineering challenges required to realise this success [19], including understanding the properties of plasma turbulence.

## Chapter 2

# Tokamak confinement

In the previous chapter it was seen that while the tokamak configuration provides the basis for a magnetic confinement device in principle, their historical performance has not demonstrated sufficiently good confinement for the purposes of energy generation. Several topics central to the confinement of tokamaks are discussed in this chapter, including the mechanisms by which radial transport can occur, different confinement regimes, and the modelling paradigms available for interfacing with tokamak experiments.

### 2.1 Collisional transport

A baseline level of radial transport occurs due to collisions in the plasma, owing to the existence of the pressure gradient across the minor radius of the tokamak. The effect of these collisions is for particles and energy to be transported down the pressure gradient, hence radially outwards, reducing confinement. This type of transport is ubiquitous in the presence of a pressure gradient and so represents an irreducible level of transport known as ‘classical’ transport.

By approximating classical transport as a diffusive process a diffusion coefficient  $D_{C1}$  may be estimated using random walk arguments, for which one can show  $D_{C1} \sim 10^{-5} \text{ m}^2\text{s}^{-1}$  [6]. This result is orders of magnitudes smaller than experimentally measured values of transport coefficients, which typically observe  $D_{\text{exp}} \sim 1 \text{ m}^2\text{s}^{-1}$  [8]. While omnipresent, this source of radial transport is clearly not responsible for the confinement properties seen in experiment. Including the effects of toroidal geometry in the analysis of collisional transport gives rise to ‘neoclassical’ transport [6, 8]. This estimates a much larger level of diffusion,  $D_{\text{Neo}} \sim 100D_{C1}$ , however is still not enough to meet with the experimentally observed levels. This discrepancy gives rise to the term ‘anomalous’ transport, accounting for the difference between

that observed in experiment and that predicted by the combination of classical and neoclassical transport.

## 2.2 Experimental character of anomalous transport

A ubiquitous feature of tokamaks is the presence of fluctuations in the plasma, which have been shown to correlate with the levels of transport observed [20, 21]. Measuring density fluctuations in the tokamak core shows the fluctuations to be small amplitude, typically  $\delta n/n_0 \sim 1\%$  [22], with heavy ion beam measurements drawing a similar conclusion for the potential fluctuations,  $e\delta\phi/T \sim \delta n/n_0$  [8]. These fluctuations can rise in the edge region, up to as much as 50%.

The frequencies of the fluctuations are observed to be centered around  $\sim 100\text{kHz} \ll \Omega_i$  [23], with a broadband signal for a given wavenumber indicative of a nonlinear process. The fluctuations are characterised by disparate correlation lengths parallel and perpendicular to the equilibrium magnetic field. Wavenumbers perpendicular to the field  $k_\perp$  typically have values on the order of the ion gyroradius  $k_\perp\rho_i \sim 1$ , peaking around 0.3; however can also exhibit electron scale fluctuations,  $k_\perp\rho_e \sim 1$ . Parallel to the field however the turbulent structures are far more elongated, such that they are comparable to the machine size,  $k_\parallel a \sim 1$ . These experimental observations provide physically-motivated orderings which can be used in the development of theories describing plasma fluctuations.

## 2.3 Turbulent transport

It is now generally accepted that the origin of anomalous transport observed in tokamaks is due to turbulence, which is associated with fluctuations within the plasma. The steep gradients in density and temperature serve as a source of free energy, which seed microinstabilities<sup>1</sup> in the presence of perturbations. The amplitudes of these perturbative oscillations grow, during which the different modes interact nonlinearly, and eventually *saturate*, providing a steady level of fluctuation. These turbulent fluctuations drive the radial transport responsible for the poor levels of confinement seen experimentally.

---

<sup>1</sup>These are named due to the correlation lengths perpendicular to the field being of the order of the ion gyroradius, in contrast to the large-scale instabilities associated with MHD being on the order of the machine size  $a$ .

### 2.3.1 Electron drift waves

#### Stable electron drift waves

A simplified model can be used to explore some of the properties of the instabilities that generate plasma turbulence. These are called drift wave instabilities [24, 25], the simplest of which is the electron drift wave.

We start in Cartesian geometry, with a uniform magnetic field aligned with the  $z$ -axis,  $\mathbf{B} = B\hat{\mathbf{k}}$ . Present are electrons and a single hydrogenic ion species, with equilibrium densities  $n_{e,0} = n_{i,0} = n_0$ . There is a uniform equilibrium density gradient in the  $-x$  direction,  $\nabla n_0 = -|dn_0/dx|\hat{\mathbf{i}}$ . We model each species as a fluid, and assume the adiabatic response for the electrons. As a closure relation to the fluid system we assume that the ions are cold,  $T_i \rightarrow 0$ .

Initially in equilibrium, we introduce a perturbation to the system described via wavevector  $\mathbf{k}$ , such that dynamical quantities contain an equilibrium and perturbed part, e.g.  $n_e = n_{e,0}(x) + \delta n_e$ , which we then linearise.

From quasineutrality, in the limit  $|\mathbf{k}|\lambda_D \ll 1$ , one has<sup>2</sup>

$$\delta n_i = \delta n_e \quad (2.1)$$

where the electron density is given by the adiabatic response

$$\delta n_e = n_0 \frac{e}{T_e} \delta \phi. \quad (2.2)$$

We thus drop the species label for the density fluctuations. The linearised continuity equation for the ions is

$$\frac{\partial \delta n}{\partial t} + (\delta \mathbf{u}_i \cdot \nabla) n_0 + n_0 (\nabla \cdot \delta \mathbf{u}_i) = 0 \quad (2.3)$$

and the linearised ion momentum equation is

$$\frac{1}{\Omega_i} \frac{\partial \delta \mathbf{u}_i}{\partial t} = \frac{1}{B} (-\nabla \delta \phi + \delta \mathbf{u}_i \times \mathbf{B}). \quad (2.4)$$

Assuming that the frequency of the turbulent fluctuations is small relative to the ion gyrofrequency,  $\omega \ll \Omega_i$ , the left-hand side can be considered negligible relative to the other terms. The momentum equation therefore becomes  $\nabla \delta \phi = \delta \mathbf{u}_i \times \mathbf{B}$ , which has the solution  $\delta \mathbf{u}_i = \delta u_{i,\parallel} \hat{\mathbf{b}} + \mathbf{v}_E$ , where  $\mathbf{v}_E = \mathbf{B} \times \nabla_{\perp} \delta \phi / B^2$  is the  $E \times B$  velocity.

Inserting this result into the ion continuity equation, it can be shown that the divergence of the  $E \times B$  velocity is zero in a uniform magnetic field, and the divergence of the parallel

---

<sup>2</sup>The validity of the ordering assumptions employed in this model will be demonstrated in Section 3.1.2.

velocity is small in the limit of small parallel wavenumber  $k_{\parallel}$ . The linearised version of the ion continuity equation is therefore

$$\frac{\partial \delta n}{\partial t} + \mathbf{v}_E \cdot \nabla n_0 = 0. \quad (2.5)$$

Using a Fourier representation for fluctuating quantities, such that for density one has

$$\delta n = \sum_{k_y, \omega} \delta \hat{n}_{k_y, \omega} e^{i(k_y y - \omega t)} \quad (2.6)$$

equation 2.5 becomes

$$\sum_{k_y, \omega} \left[ -i\omega \delta \hat{n}_{k_y, \omega} + ik_y \frac{1}{B} \left| \frac{dn_0}{dx} \right| \delta \hat{\phi}_{k_y, \omega} \right] e^{i(k_y y - \omega t)} = 0. \quad (2.7)$$

Substituting equation 2.2, this yields the dispersion relation

$$\omega_0(k_y) = \frac{T_e}{eBL_n} k_y \quad (2.8)$$

where  $L_n = n_0/|dn_0/dx|$  is the length scale of the density gradient. We note that the frequency  $\omega_0$  is purely real, and so there is no growth or suppression of the fluctuations. It is also positive, with the resulting wave travelling in the  $+y$  direction.

A diagram of the mechanism of the electron drift wave is shown in figure 2.1 for a single Fourier mode, taking a perturbation of the density of the form  $\delta n = A_n \sin(k_y y - \omega_0 t)$  for fluctuation amplitude  $A_n$ . This change in density causes an instantaneous change in the electrostatic potential through the adiabatic response of the electrons,  $\delta \phi = A_\phi \sin(k_y y - \omega_0 t)$ , which generates a fluctuating electric field,  $\delta \mathbf{E} = -\nabla \delta \phi = -A_\phi k_y \cos(k_y y - \omega_0 t) \hat{\mathbf{j}}$  (not shown). This in turn generates a fluctuating  $E \times B$  velocity,  $\delta \mathbf{v}_E = -(A_\phi k_y / B) \cos(k_y y - \omega_0 t) \hat{\mathbf{i}}$ , which at all points along the perturbation can be seen to convect the plasma back towards its equilibrium position without instability.

This model is excessively simple for application to tokamaks, however it has demonstrated that in the presence of an equilibrium density gradient and an equilibrium magnetic field, a fluctuation can be supported. In this case, the adiabatic response of the electrons ensures that the density fluctuations are always in phase with the potential, which causes the resulting wave to be stable. By introducing a phase shift between these fluctuations an instability can be driven, as shall now be shown.

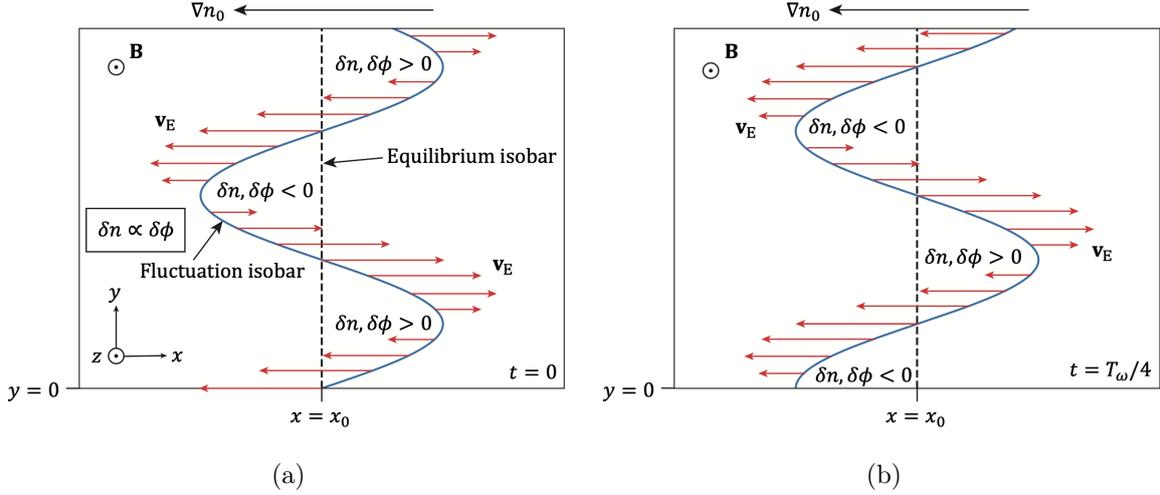


Figure 2.1: The mechanism of the electron drift wave, at some initial time  $t = 0$  (a) and a quarter of a period later,  $t = T_\omega/4$  (b). Note the movement of the wave in the positive  $y$  direction between (a) and (b).

### Unstable electron drift waves

A brief discussion on the relationship between the imaginary part of a Fourier mode amplitude and the phase of its oscillation, including in the case of two quantities oscillating with the same frequency, is given in appendix A.3. Relaxing the adiabatic response approximation, we consider some physical process by which a *non-adiabatic* contribution  $\delta n_{\text{NA}}$  to the density fluctuations is induced, such as the effect of collisions or finite electron inertia [25]. This replaces equation 2.2 with

$$\delta n = \frac{en_0}{T_e} \delta\phi + \delta n_{\text{NA}} \quad (2.9)$$

where the Fourier amplitudes are related by

$$\begin{aligned} \delta \hat{n}_{k_y, \omega} &= \frac{en_0}{T_e} \delta \hat{\phi}_{k_y, \omega} + \delta \hat{n}_{\text{NA}, k_y, \omega} \\ &= \frac{en_0}{T_e} \delta \hat{\phi}_{k_y, \omega} Z_\Delta \end{aligned} \quad (2.10)$$

where  $Z_\Delta = \left(1 + \delta \hat{n}_{\text{NA}, k_y, \omega} / \left(en_0 \delta \hat{\phi}_{k_y, \omega} / T_e\right)\right)$ . It then follows

$$|\delta \hat{n}_{k_y, \omega}| e^{i\theta_n} = \frac{en_0}{T_e} \left| \delta \hat{\phi}_{k_y, \omega} \right| \left| Z_\Delta \right| e^{i(\theta_\phi + \theta_\Delta)}. \quad (2.11)$$

where  $\theta_\phi$  and  $\theta_\Delta$  are the phases of the potential fluctuations and  $Z_\Delta$  respectively. The phase of the density fluctuation relative to the potential is therefore  $\theta_n = \theta_\phi + \theta_\Delta$ , with a phase difference present given  $\theta_\Delta \neq 0$ .

Repeating the analysis of the drift wave, now with equation 2.9 replacing 2.2, one finds a dispersion relation which includes both a real part to the frequency of the mode  $\tilde{\omega}_{k_y}$  and an

imaginary part  $\gamma_{k_y}$ ,

$$\begin{aligned}\omega_{k_y} &= \tilde{\omega}_{k_y} + i\gamma_{k_y} \\ &= \omega_0 \frac{\text{Re}(Z_\Delta^*)}{|Z_\Delta|^2} + i\omega_0 \frac{\text{Im}(Z_\Delta^*)}{|Z_\Delta|^2}.\end{aligned}\quad (2.12)$$

Here we see the imaginary part of the dispersion relation originates from a phase shift between the potential and density fluctuations,  $\theta_\Delta$ . For  $\text{Im}(Z_\Delta^*) > 0$  this corresponds to exponential growth of the mode, and hence an instability. For the density fluctuations, taking  $\theta_\phi = -\pi/2$  one has

$$\begin{aligned}\delta n &= \delta n_{k_y} e^{\gamma_{k_y} t} e^{i(k_y y - \tilde{\omega}_{k_y} t)} + \delta n_{k_y}^* e^{\gamma_{k_y} t} e^{-i(k_y y - \tilde{\omega}_{k_y} t)} \\ &= A_n e^{\gamma_{k_y} t} \sin(k_y y - \tilde{\omega}_{k_y} t + \theta_\Delta)\end{aligned}\quad (2.13)$$

where we have used the parity property of the frequency, which can be shown to be  $\omega_{-k_y} = -\tilde{\omega}_{k_y} + i\gamma_{k_y}$ . Correspondingly for the potentials,

$$\delta\phi = A_\phi e^{\gamma_{k_y} t} \sin(k_y y - \tilde{\omega}_{k_y} t)\quad (2.14)$$

and hence both quantities grow with growth rate  $\gamma_{k_y}$ . The case for maximum instability,  $\theta_\Delta = -\pi/2$ , is shown in figure 2.2, corresponding to the potential lagging behind the density by a quarter period. In this case, the resulting fluctuating  $E \times B$  velocity is now positioned such that the point of maximum velocity occurs at the position of maximum displacement, reinforcing the perturbation and causing an instability.

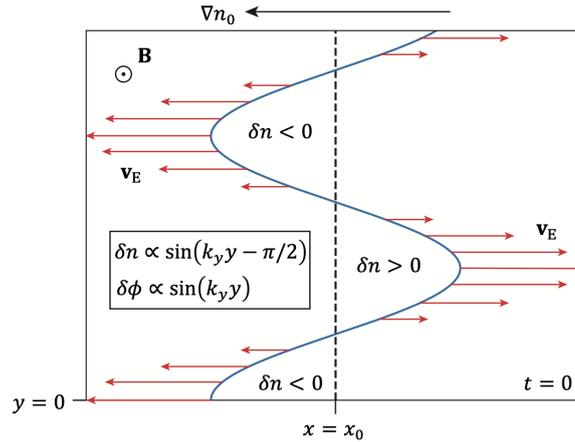


Figure 2.2: A drift wave now with a phase shift of  $\pi/2$  between the density and the potential. The resulting  $E \times B$  motion now acts to further perturb the density, resulting in an instability.

This analysis of the electron drift wave has introduced some key concepts regarding the dynamics of microinstabilities which can generate turbulence, however for simplicity we have

left out a number of relevant effects, including general geometry, a kinetic treatment and collisions. The inclusion of these effects can illuminate a number of other low frequency ( $\omega \ll \Omega_i$ ) instability modes [25], the most important of which for core transport in large aspect ratio tokamaks shall now be summarised.

### 2.3.2 Other drift waves

#### Ion temperature gradient mode (ITG)

The ITG is an ion-scale instability ( $k_{\perp}\rho_i \sim 1$ ) which is thought to drive a great deal of the anomalous transport in the ion heat channel in modern tokamaks. The instability mechanism is similar to that of the electron drift wave, however relates to a perturbation in the ion temperature  $\delta T_i$  in the presence of a temperature gradient rather than the density gradient. For a perturbation in ion temperature the magnetic drift velocities (equation 1.47) of the ions originating from the hotter region are greater than those from the colder region. This disparate motion of the ions causes a charge separation and hence an  $E \times B$  drift, setting up an unstable drift wave, as the temperature and the potential do not oscillate in phase by this mechanism.

The instability of the ITG is relevant to the outboard side of the tokamak, due to the ‘bad curvature’ induced by the orientation of the curvature vector of the magnetic field relative to the pressure gradient<sup>3</sup>. On the inboard side of the tokamak this orientation is reversed, with the resulting ‘good curvature’ causing the feedback mechanism to be inverted and thus suppress the instability.

#### Trapped electron mode (TEM)

The TEM is an ion-scale drift wave instability driven by a pressure gradient in the presence of ‘trapped electrons’. These are a subset of the electron population confined to the outboard side of the tokamak as a result of charged particle motion in a spatially-varying magnetic field [6]. The remaining electrons are known as ‘passing electrons’, which are able to traverse the entire poloidal extent of the tokamak.

The inability of a fraction of the electron population to complete a poloidal transit is a source of phase difference between the density perturbations and the potential, and thus in the presence of a perturbation an instability is driven. Due to the physics of this mode being centred around the dynamics of the electrons it cannot be described under the assumption of

---

<sup>3</sup>This is applicable to the ‘curvature driven’ ITG, a ‘slab’ ITG can be supported in the absence of curvature.

the adiabatic electron response.

### **Electron temperature gradient mode (ETG)**

The ETG is the analogy of the ITG for electrons, and thus is relevant at a wavelength comparable with the electron gyroradius,  $k_{\perp}\rho_e \sim 1$ . This small wavelength causes the ETG to only provide a relatively small degree of transport, however can couple to other, longer wavelength modes. This mode also cannot be captured under the assumption of adiabatic electrons.

## **2.4 Empirical scaling laws and modes of tokamak operation**

The historical lack of understanding of the source of anomalous transport, as well as the difficulty in the modelling of transport due to plasma turbulence, necessitated the use of 0D scaling laws to understand how the confinement of tokamaks is affected by different plasma parameters. By collating results over many machines, one can use fitting and regressions to produce relations between the confinement time  $\tau_E$  and relevant parameters. The powers of the parameters to be fitted are the plasma current  $I$ , the applied heating power  $P$ , the toroidal magnetic field strength  $B_t$ , the plasma density  $n$ , the average atomic mass  $A$ , the major radius  $R_0$ , the inverse aspect ratio  $\bar{\epsilon} = a/R_0$ , and the elongation of the poloidal cross-section  $\kappa$ .

While turbulent transport has been observed to be a ubiquitous feature of tokamak experiments, it has been seen that different levels of transport can be observed under different experimental conditions, known as confinement modes, of which the most pertinent shall now be introduced. The different confinement modes generally behave differently as a function of these parameters, and scaling laws must be made for each confinement mode separately.

### **Ohmic plasmas**

In the initial stages of a plasma pulse, the toroidal current that is driven experiences resistance due to electron-ion collisions, which causes the temperature of the plasma to rise. This process is known as Ohmic heating. The amount of heating that can be obtained in this way however is fundamentally limited, as the electron-ion collision frequency can be shown to scale inversely with the electron temperature  $\nu_{ei} \propto T_e^{-3/2}$ . Therefore as the plasma gets hotter, the collision frequency and hence resistivity decreases, reducing its ability to generate its own heating. The maximum temperature reached by this method is not enough to produce significant fusion [8].

Within the regime of ohmically heated plasmas, two further delineations can be made, due to the experimentally-observed dependence on the confinement time with average density [26]. At low densities, the confinement time scales linearly,  $\tau_E \propto n$ , known as the ‘linear Ohmic confinement’ (LOC) regime. This proportionality with density applies up until a point at which the confinement time reaches a maximum, and will no longer rise with increasing density. This is known as the ‘saturated Ohmic confinement’ (SOC) regime.

### L-modes

Once the limitation of temperature gain due to Ohmic heating has been reached, alternative methods can be used to continue raising the plasma temperature. These methods are known as auxiliary heating.

Such methods include neutral beam injection (NBI), in which high-energy atoms are fired into the plasma. Via collisions, these atoms become ionised and thermalise with the plasma, imparting their energy and raising its temperature. Heating via electromagnetic waves can also be used, via radiofrequency heating, as well as cyclotron resonance heating. This technique comprises tuning the frequency of the radiation to the gyrofrequency of the electrons/ions in a given region of the magnetic field, exploiting a resonance in the frequency for an efficient transfer of energy, allowing for localised heat deposition. Generally a combination of these methods can be used to raise the plasma to fusion temperatures. A scaling law for the confinement time used for L-modes is  $\tau_{E,\text{th}}^L$ , given by [27]

$$\tau_{E,\text{th}}^L = 0.058 I^{0.96} B_t^{0.03} P^{-0.73} n^{0.40} A^{0.20} R_0^{1.83} \bar{\epsilon}^{-0.06} \kappa^{0.64}. \quad (2.15)$$

One observes a profoundly unfavourable scaling of confinement time with power injection  $P$ , with the confinement time scaling almost *inversely* with the applied heating power. This indicates that as one attempts to heat the plasma, it further expels energy, due to the heating acting a source of free energy for the turbulent fluctuations.

### H-modes

During an experimental campaign on the tokamak ASDEX, it was found that with a sufficient amount of heating, the plasma spontaneously reached a state of improved confinement [28]. This state is called the H-mode (for ‘high’ confinement), rendering plasmas heated with auxiliary heating outside the H-mode as L-modes (‘low’ confinement).

The H-mode is characterised by a sudden decrease in radial transport, particularly in the edge region, due to the formation of an ‘edge transport barrier’. This reduction in transport leads

to the formation of a ‘pedestal’ in the edge region, where the enhanced confinement causes the radial profiles of density and temperature to rise sharply. The profiles in the core then ‘sit atop’ this pedestal, increasing the energy density of the plasma across the minor radius. H-mode operation has now been obtained in most tokamaks, and an empirical blueprint exists for reaching this higher state of confinement (the ‘L-H transition’). While the reduced radial transport of H-modes is advantageous for plasma confinement, they are however subject to transient MHD instabilities called edge-localised modes (ELMs). These are characterised by expulsions of energy and particles at the plasma boundary, which can damage the surrounding vessel. The understanding and control of ELMs, as well as the theoretical understanding of the L-H transition, continue to be active areas of research [29, 30].

A commonly used scaling law in the presence of ELMy H modes is IPB98( $y, 2$ ), given by [18]

$$\tau_E^{\text{IPB98}(y,2)} = 0.145 I^{0.93} B_t^{0.15} P^{-0.69} n^{0.41} A^{0.19} R_0^{1.97} \epsilon^{0.58} \kappa^{0.78}. \quad (2.16)$$

In the case of ELM-free operation, one has [8]

$$\tau_E^{\text{ELM free}} = 0.068 I^{0.94} B_t^{0.27} P^{-0.68} n^{0.34} A^{0.43} R_0^{1.98} \epsilon^{0.10} \kappa^{0.68}. \quad (2.17)$$

We note that the scalings presented in this section are not the only ones available for these confinement regimes. While these relations are useful for understanding how confinement varies with tokamak parameters and interpolating to different regimes, they are very limited in their ability to extrapolate to unexplored regimes, and provide little to no insight as to the causes behind the relations.

## 2.5 First principles modelling

To partially alleviate reliance on confinement scaling laws and their associated issues, efforts have been made to be able to estimate confinement from so-called first principles. Instead of the zero-dimensional considerations of the previous section, predictions of the radial profiles of the plasma can be obtained using physics-based models.

Through the use of the fluid equations discussed in section 1.2.1, one can obtain transport equations which describe the large scale evolution of the macroscopic plasma quantities, namely the density, momentum and temperature of each species. By averaging these equations over the flux surfaces, they describe the 1D evolution of these quantities across the minor radius with time. For example, the density transport equation is [31]

$$\frac{1}{V'} \frac{\partial}{\partial t} (\langle n_s \rangle_{\text{FS}} V') + \frac{1}{V'} \frac{\partial}{\partial \varrho} (V' \Gamma_s) = \langle S_{n_s} \rangle_{\text{FS}} \quad (2.18)$$

where  $V' = \partial V / \partial \varrho$  is the radial derivative of the volume enclosed by the surface<sup>4</sup>,  $\Gamma_s$  is the particle flux of species  $s$ ,  $S_{n_s}$  is the particle source of species  $s$  and  $\langle \dots \rangle_{\text{FS}}$  denotes an average over the flux surface. Similar equations can be derived for the evolution of momentum and energy, with the equations of tokamak force balance and the evolution of the equilibrium magnetic field through Faraday's law completing the system.

These transport equations are affected by almost every physical consideration in the tokamak, a (non-exhaustive) list of which includes tokamak geometry, the transport of impurities and neutrals, auxiliary heating, MHD instabilities, as well as classical, neoclassical and turbulent transport. This approach therefore requires each aspect of the tokamak to be simulated together, such that they can interface with each other and play a role in their respective evolution. For a given set of initial conditions the system is evolved forwards explicitly in time, simulating the global evolution of the plasma on confinement timescales.

Naturally this system of equations offers greater opportunity for local extrapolation in parameter space than scaling laws, provided the physics models are suitably general. This first principles approach is called *integrated modelling*, examples of which include the codes JETTO [32] and ASTRA [33]. A central part of integrated modelling is the calculation of turbulent fluxes, as this forms the dominant transport mechanism in tokamaks. Models by which this can be calculated will now be considered.

## 2.5.1 Turbulent transport models

### Nonlinear gyrokinetics

The most accurate paradigm available for simulating tokamak turbulence is that of *nonlinear gyrokinetics* [34–38]. This is a kinetic description similar to that of the 6D equation given by 1.18, however by exploiting the rapidity of the gyromotion relative to other timescales in the turbulent system the gyrophase dependence can be integrated over, reducing the dimensions of the problem to 5. Being a kinetic description and thus not required to provide closure to the system, this represents the most accurate theoretical framework for modelling plasma turbulence, and has been extensively validated against experiment [39–46].

Gyrokinetics can be simulated both *globally*, in which the turbulence is modelled over a sufficient extent of the radial domain to include variation in equilibrium quantities, and *locally*, which focuses around a field line on a single flux surface to determine the turbulent properties

---

<sup>4</sup>For simplicity, the form of equation 2.18 assumes that the flux-surface label denoted by  $\varrho$  does not vary in time.

of a given local equilibrium. For integrated modelling purposes, one could conceive of setting up a 1D grid across the minor radius, evaluating the turbulence locally at each flux surface and building a radial profile of the turbulent transport.

However, the great computational expense of even local nonlinear gyrokinetics renders it impractical for routine use in integrated modelling simulations<sup>5</sup>, with the resources required for a single local simulation typically being at least on the order of  $10^4 - 10^5$  CPUh [49], and potentially being far larger. For reasonable estimates of a grid-size of 30 radial points and a time step of  $\Delta t \sim 10^{-3}$ , the simulation of 1 second of plasma evolution would require  $\sim 10^4$  nonlinear gyrokinetic simulations, necessitating alternative approaches to turbulent flux calculation.

### Quasilinear transport models

To enable calculations of turbulent transport on a timescale commensurate with integrated modelling a class of model known as a *quasilinear* transport model is used, which through the use of various simplifications allows one to obtain relatively fast estimates of the turbulent transport for a given local equilibrium. Examples of quasilinear models include TGLF [50] and QuaLiKiz [49]. These models bypass the expense of calculating the fluxes in the nonlinear gyrokinetic system by instead solving for the linear response of the plasma instabilities, which is then combined with an estimation of the magnitude of the saturated turbulent fluctuations via a *saturation rule* to provide a calculation of the fluxes in a greatly reduced time. Quasilinear models have been extensively validated against nonlinear gyrokinetic codes [51, 52] and have been used in integrated modelling suites to successfully simulate experimental plasma discharges [53–57].

The simplifications required for quasilinear transport models can naturally incur error in the predicted turbulent fluxes compared to nonlinear gyrokinetic results. They are typically trained against nonlinear gyrokinetic simulations, which can cause them to perform less reliably in parameter spaces outside of those they have been trained on. This subjects quasilinear models to the requirement of constant development and verification as new parameter regimes are explored. A cycle of such development is of the form<sup>6</sup>:

---

<sup>5</sup>There are models for which one can solve the 1D transport equations to steady state using nonlinear gyrokinetic codes, such as TRINITY [47] and TGYRO [48], however these cannot describe transient phenomena on confinement timescales.

<sup>6</sup>This cycle assumes that local nonlinear gyrokinetics forms the ground truth for the description of plasma turbulence. In regimes for which the gyrokinetic ordering assumptions (section 3.1.2) are not well-satisfied, one could include an additional step comparing nonlinear gyrokinetics to experiment where necessary.

1. Identify failure of quasilinear models
2. Compare contemporary quasilinear models to nonlinear gyrokinetic results
3. Identify systematic discrepancies, and diagnose why these are present
4. Improve the quasilinear models through understanding of nonlinear gyrokinetic results
5. Validate integrated modelling simulations with new quasilinear model against experiment

By continuing to develop quasilinear models, their reliability in already-explored regimes increases, and their ability to explore extrapolated regimes is improved. An as-yet relatively unexplored parameter which is of timely importance is the ion mass  $A$ , which will now be discussed.

## 2.6 The isotope effect

With the advent of ITER comes a change in the principal fuel used in fusion experiments, from the historically ubiquitous D-D reactions to D-T. Such a change raises the important question of how the confinement properties of the plasma change with the isotope mass, in particular the levels of turbulent transport.

### GyroBohm estimates

A theoretical estimate of the dependence of the transport on isotope mass can be constructed using dimensional arguments. Approximating turbulent transport as a diffusive process, one can define a diffusion coefficient  $D_{\text{turb}} \sim \omega/k_{\perp}^2$  for characteristic frequency  $\omega$  and perpendicular wavenumber  $k_{\perp}$  of the fluctuations. Assuming  $k_{\perp} \sim 1/\rho_i$  from section 2.2 and using equation 2.8 to estimate  $\omega \sim T/(eBa\rho_i) \sim v_{\text{th},i}/a \sim (\rho_i/a)\Omega_i$ , we obtain  $D_{\text{turb}} \sim (\rho_i/a)\rho_i^2\Omega_i \propto \sqrt{A}$ . Estimating the turbulent particle flux  $\Gamma$  via  $\Gamma \sim D_{\text{turb}}\nabla n$  one finds  $\Gamma \propto \sqrt{A}$ , and similarly for the heat flux  $Q \propto \sqrt{A}$ . These are the gyroBohm<sup>7</sup> estimates of the fluxes, which suggest that turbulent transport increases with the square root of the isotope mass.

A simple estimate of the confinement time  $\tau_E$  can then be made by approximating the time it takes for the plasma to diffuse across an area of  $\sim a^2$ , the cross-sectional area of the

---

<sup>7</sup>The name gyroBohm originates from the factor of the normalised gyroradius  $\rho_i/a$  multiplying the previously established Bohm scaling,  $D_{\text{Bohm}} = \rho_i^2\Omega_i$  [58].

tokamak. This gives  $\tau_E \sim a^2/D_{\text{turb}} \propto 1/\sqrt{A}$ . GyroBohm arguments therefore suggest that confinement should *worsen* as the isotope mass  $A$  is increased due to the greater levels of turbulent transport, which has negative implications for the use of D-T over D-D.

### Experimental scalings

While tritium discharges have been historically relatively rare, discharges of H plasmas have been widely conducted. These can be used to measure how the confinement properties of the plasma vary with isotope mass relative to D, allowing for  $A$  dependence to be determined for the confinement law scalings of section 2.4. Examining the confinement scalings of equations 2.15-2.17, it is seen that none of them exhibit a dependence in agreement with gyroBohm arguments,  $\tau_{E,\text{GB}} \propto A^{-0.5}$ . For all of the scaling laws, instead a *positive* exponent is found, implying experimentally that the confinement time *increases* with ion mass, in direct opposition to that suggested by gyroBohm scaling. This discrepancy between the experimental dependence of  $\tau_E$  with  $A$  and that predicted by gyroBohm arguments is called ‘the isotope effect’ [59–62]. Whilst this result bodes well for the success of future tokamak devices in their shift from pure D plasmas to the operational mix of D-T, the theoretical understanding for this discrepancy remains incompletely understood.

### The isotope effect in gyrokinetics

To investigate the gyroBohm scaling assumption of fluxes  $\Gamma_{\text{GB}}, Q_{\text{GB}} \propto \sqrt{A}$ , nonlinear gyrokinetic codes can be used to accurately simulate how local turbulent fluxes depend on isotope mass. Previous studies have shown that in sufficiently simple cases, namely those dominated by ITG turbulence with a single ion species and adiabatic electrons, the gyroBohm scaling of fluxes is produced [54, 63]. The inclusion of more sophisticated physics in these simulations breaks this scaling, as has been demonstrated for the inclusion of kinetic electrons [64], electromagnetic effects [65], and collisions [66, 67]. Numerous studies have also shown that not only can local fluxes deviate from the gyroBohm prediction but can follow the opposite trend, the so-called *anti-gyroBohm* scaling, in which fluxes scale inversely with isotope mass  $Q \propto A^p$ ,  $p < 0$  [64, 68]. The plurality of mechanisms involved in this gyroBohm-breaking effect makes even a partial explanation of the isotope effect on a local level challenging.

### The isotope effect in quasilinear models

While quasilinear models have been used successfully in the modelling of deuterium discharges, their historically limited considerations of plasmas in other isotopes, coupled with their in-

complete description of the isotope effect, cause them to struggle to replicate the behaviour of nonlinear gyrokinetic fluxes with isotope mass. This renders predictive flux-driven integrated modelling of non-deuterium plasmas to be unreliable [54, 68], undermining prediction and optimisation efforts of experimental campaigns.

In view of D-T operations in ITER and other future machines it is therefore essential that modern quasilinear models more accurately approximate fluxes in plasmas of different isotopes. Their current inability to do so represents a failure of the models, as outlined in section 2.5.1, and so constitutes an opportunity to execute a cycle of quasilinear model development in this area.

## 2.7 Topics addressed in this thesis

These first two chapters have now discussed all introductory topics needed to understand the purpose of this thesis, which shall now be summarised.

In a tokamak plasma, turbulent transport forms the dominant mechanism by which the confinement of particles and energy is lost. To make predictions regarding the performance of tokamak experiments integrated modelling suites can be employed, with the turbulent transport being calculated via quasilinear models. The simplifications made to these models that allow for their necessary computational efficiency however can make them less reliable outside of the parameter space in which they have been trained. Where this is the case, the models require amending through a cycle of quasilinear model development.

A parameter which has seen relatively little validation in quasilinear models is that of isotope mass  $A$ . Tokamak experiments demonstrate the isotope effect, in which the scaling of the confinement time  $\tau_E$  with  $A$  is found to be in opposition to that suggested by simple gyroBohm arguments. The non-trivial nature of the turbulent behaviour and hence confinement time with isotope mass severely weakens the reliability of quasilinear models in integrated modelling simulations of non-deuterium plasmas. In anticipation of D-T operations in future devices this is an essential issue to address, which is what this work seeks to achieve.

By comparing contemporary quasilinear models to nonlinear gyrokinetic simulations, discrepancies between the isotope dependence of transport will be demonstrated explicitly, indicating the need for additional physics in their description of the turbulence. The cause of this will then be diagnosed and parameterised, and used to develop improvements to the quasilinear models to better match nonlinear gyrokinetics. The new quasilinear model will then be validated in integrated modelling simulations comparing against experimental data, including

the JET isotope experiments in H, D and T conducted over the course of this PhD as part of the DTE-2 campaign [19].

The remainder of this thesis is formulated as follows. Chapter 3 derives the gyrokinetic theory, which forms the principal theoretical framework for turbulent plasma dynamics used in this work. Chapter 4 then discusses the structure of quasilinear turbulence models, which is followed by an analysis of the discrepancies between current quasilinear models and nonlinear gyrokinetic results. The development of a new quasilinear saturation rule SAT3 is detailed in Chapter 5, addressing the discrepancies observed in the previous chapter. Chapter 6 then showcases results obtained from using SAT3 in integrated modelling simulations against experimental data obtained from JET, with comparison to previous saturation models. Finally, Chapter 7 concludes the work, summarising its most important points and outlining ideas for future study.

# Chapter 3

## Gyrokinetic theory

### 3.1 Introduction

In this chapter, the theoretical framework with which to describe the dynamics of turbulent plasmas is discussed. Building on the single particle dynamics of the constant field case presented in Section 1.2.4, the equations of motion for a charged particle in turbulent fields are first obtained, which are then used in the derivation of the gyrokinetic system. An analysis of coordinate systems useful for the description of tokamak geometry follows, which culminates in a discussion on the local limit of gyrokinetics.

We start by introducing two concepts of the gyrokinetic theory which are essential in the ensuing derivations, namely the ensemble average and the gyrokinetic ordering assumptions.

#### 3.1.1 The ensemble average

The ensemble average is a fundamental operation to a turbulence theory. It is the process by which one can separate a given quantity into its equilibrium and fluctuating parts, allowing the evolution of each to be considered separately. Conceptually it is more of a theoretical average than a strict mathematical operation, as it involves considering infinitely many repetitions of the same ‘experiment’. That is, one conducts said experiment involving a turbulent quantity  $f(\mathbf{r}, t)$  and measures its value at a position and time, providing the first measurement  $f^{(1)}(\mathbf{r}, t)$ . Then one considers conducting the same experiment again, and measuring the same quantity at the same position and time,  $f^{(2)}(\mathbf{r}, t)$ . As a consequence of the turbulent fluctuations, the measured value will generally be different than the first measurement,  $f^{(1)}(\mathbf{r}, t) \neq f^{(2)}(\mathbf{r}, t)$ <sup>1</sup>. Repeating this process over a large number of repetitions and taking

---

<sup>1</sup>This can be considered analogous to the exercise of repeatedly rolling a die.

the average over the set defines the ensemble average,

$$\langle f \rangle_{\text{Ens}}(\mathbf{r}, t) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f^{(i)}(\mathbf{r}, t) \quad (3.1)$$

which we see is essentially an expectation value of a probability distribution. One can therefore decompose a quantity  $f(\mathbf{r}, t)$  via

$$f(\mathbf{r}, t) = \langle f \rangle_{\text{Ens}}(\mathbf{r}, t) + \delta f(\mathbf{r}, t) \quad (3.2)$$

where  $\delta f$  is the fluctuating part of the function  $f$ . Using equation 3.1, some properties of the ensemble average are, for two functions  $f$  and  $g$ :

- Linearity:  $\langle f + g \rangle_{\text{Ens}} = \langle f \rangle_{\text{Ens}} + \langle g \rangle_{\text{Ens}}$
- Effect of multiple averages:  $\langle \delta f \rangle_{\text{Ens}} = 0$ ,  $\langle \langle f \rangle_{\text{Ens}} \rangle_{\text{Ens}} = \langle f \rangle_{\text{Ens}}$
- Commutation with derivatives:  $\langle \partial f / \partial z^i \rangle_{\text{Ens}} = \partial \langle f \rangle_{\text{Ens}} / \partial z^i$ , for some phase coordinate  $z^i$
- Product rule:  $\langle fg \rangle_{\text{Ens}} = \langle f \rangle_{\text{Ens}} \langle g \rangle_{\text{Ens}} + \langle \delta f \delta g \rangle_{\text{Ens}}$

The last term of the product rule encapsulates two fluctuating parts  $\delta f$  and  $\delta g$ , both varying on the scale of the turbulence, manifesting a contribution to the ensemble-averaged scale. This shall be seen later to be the way in which turbulent fluctuations can produce transport on the macroscopic scale in tokamaks.

The ensemble average can be considered similar to averaging over length and time scales that are large compared with the turbulence and short compared to the macroscopic scale, however it is not generally the same. The ensemble average constitutes taking the average over the set of experiments at the same time and position, whereas averaging over time and space is performed for the same experimental instance<sup>2</sup>. The reader is recommended [69] for further discussion of the difference between ensemble averaging and averaging over position and time, however as an example, it is shown in appendix B.1 that time averaging of the form

$$\bar{f}(t) = \frac{1}{T_0} \int_{t-T_0/2}^{t+T_0/2} f(t') dt' \quad (3.3)$$

for a signal across a domain  $\tau > T_0$  does not satisfy the multiple averages condition, i.e  $\bar{\bar{f}} \neq \bar{f}$ .

---

<sup>2</sup>In the case of a constant ensemble average it can be shown that the two become equivalent, which shall be used later in our considerations of local gyrokinetics.

## Turbulent fields

In the turbulent plasma system, the ensemble average can be used to decompose the electric and magnetic fields. Following [31], we now write the total electric and magnetic fields using a tilde,  $\mathbf{E} \rightarrow \tilde{\mathbf{E}}$ ,  $\mathbf{B} \rightarrow \tilde{\mathbf{B}}$  such that under ensemble average decomposition they become

$$\tilde{\mathbf{E}} = \mathbf{E} + \delta\mathbf{E} \quad (3.4)$$

$$\tilde{\mathbf{B}} = \mathbf{B} + \delta\mathbf{B} \quad (3.5)$$

where for brevity in the following we have now defined the ensemble averaged parts as being without tildes,  $\mathbf{E} = \langle \tilde{\mathbf{E}} \rangle_{\text{Ens}}$  and  $\mathbf{B} = \langle \tilde{\mathbf{B}} \rangle_{\text{Ens}}$ .

The ensemble average can then be applied to Maxwell's equations. This splits each of the four equations into an averaged and fluctuating part, giving a total of eight. Taking Gauss's law as an example, one obtains

$$\nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0} \rho \quad (3.6)$$

$$\nabla \cdot \delta\mathbf{E} = \frac{1}{\epsilon_0} \delta\rho. \quad (3.7)$$

Note the total charge density  $\tilde{\rho}$  has also been decomposed,  $\tilde{\rho} = \rho + \delta\rho$ , with the same applying to the current density in Ampère's law  $\tilde{\mathbf{J}} = \mathbf{J} + \delta\mathbf{J}$ . When expressing the turbulent fields in terms of the electromagnetic potentials these also become ensemble averaged, for example  $\tilde{\mathbf{B}} = \nabla \times \tilde{\mathbf{A}} \implies \mathbf{B} = \nabla \times \mathbf{A}$ ,  $\delta\mathbf{B} = \nabla \times \delta\mathbf{A}$ .

### 3.1.2 Gyrokinetic ordering assumptions

The gyrokinetic theory is a perturbation theory, and thus requires the use of ordering. That is, the relations between the magnitudes of the different physical quantities in the system are assumed, so as to be able to determine which effects are negligible and can be disregarded. To provide numerical estimates of the sizes of these quantities, approximate values will be calculated using reference values proposed for the ITER tokamak [70] using deuterium ions, with electron density  $n_e = 10^{20} \text{ m}^{-3}$ , temperature  $T_e = 10 \text{ keV}$ , minor radius  $a = 2 \text{ m}$ , major radius  $R_0 = 6 \text{ m}$  and magnetic field strength  $B = 5 \text{ T}$ . For simplicity, in this work we will assume no equilibrium electric field  $\mathbf{E} = \mathbf{0}$  and no electromagnetic fluctuations,  $\delta\mathbf{B} = \mathbf{0}^3$ . Interested readers are recommended [31] and [71] for such an analysis.

A fundamental length scale in the consideration of single particle motion in a strong magnetic field is the ion gyroradius,  $\rho_i$ , representing the scale of the gyromotion. A characteristic value

---

<sup>3</sup>We note here that non-zero magnetic fluctuations are typically described using two scalar fields  $\delta A_{\parallel}$  and  $\delta B_{\parallel}$ , as one may show in the gyrokinetic ordering  $\delta\mathbf{B} = \delta B_{\parallel} \hat{\mathbf{b}} + \nabla_{\perp} \delta A_{\parallel} \times \hat{\mathbf{b}}$ .

of this quantity for a species  $s$  can be obtained by replacing the gyromotion speed with the thermal speed,  $\rho_i = \sqrt{m_i T_i} / (q_i B) = v_{\text{th},i} / \Omega_i$ . Using our reference values we find  $\rho_i \sim 3$  mm,  $v_{\text{th},i} \sim 7 \times 10^5$  ms $^{-1}$  and  $\Omega_i \sim 2 \times 10^8$  Hz. A second length scale in the system is that of the variation of the macroscopic fields,  $B / |\nabla B|$ , which in tokamaks can be taken to be on the order of the minor radius<sup>4</sup>  $a \sim 2$  m  $\gg \rho_i$ . The difference in size between these length scales is described via the normalised gyroradius  $\rho_*$ , defined as

$$\rho_* \equiv \frac{\rho_i}{a} \quad (3.8)$$

for which we have  $\rho_* \sim 10^{-3}$ .

We now define a number of additional parameters  $\epsilon_i$  to quantify the differences in magnitude between other scales in the system, making use of our reference values and experimental observations of fluctuations described in section 2.2. For the size of the fluctuations relative to the background we introduce  $\epsilon_\delta$ , such that  $\delta n/n \sim e\delta\phi/T \sim \epsilon_\delta$ . From section 2.2, this has a value of  $\sim 10^{-3} - 10^{-2}$  in the tokamak core.

The difference in the length scales of the fluctuations parallel and perpendicular to the field is characterised by  $k_{\parallel}/k_{\perp} \sim \epsilon_{\perp}$ . Experimentally it is observed that fluctuations approximately follow  $k_{\parallel} \sim \mathcal{O}(1/a)$  and  $k_{\perp} \sim \mathcal{O}(1/\rho_i)$ , hence in the tokamak core  $\epsilon_{\perp} \sim 10^{-3}$ . Ensemble averaged quantities vary on the macroscopic length scale,  $\sim \mathcal{O}(1/a)$  both perpendicular and parallel to the field.

For the frequency of the fluctuations  $\omega_{\text{turb}}$  relative to the ion gyroradius we use  $\epsilon_\omega \sim \omega_{\text{turb}}/\Omega_i$ , for which one finds  $\epsilon_\omega \sim 100$  kHz/200 MHz  $\sim 10^{-3}$ . Ensemble averaged quantities vary in time on the confinement timescale  $\sim 1/\tau_E$ , where  $\tau_E \sim \mathcal{O}(1$  s), and thus defining  $\epsilon_\tau \sim 1/(\Omega_i \tau_E)$  we have  $\epsilon_\tau \sim 10^{-9}$ . For this reason, ensemble averaged quantities are typically referred to as ‘equilibrium’ quantities, and the two terms can be used relatively interchangeably.

We also are required to order the Debye length  $\lambda_D$  and the speed of light  $c$ , which appear in Maxwell’s equations. We thus define  $\epsilon_{\lambda_D} \sim (\lambda_D/\rho_i)^2$  and  $\epsilon_c \sim v_{\text{th},i}/c$ . Using  $c \sim 3 \times 10^8$  ms $^{-1}$  and the characteristic values to obtain  $\lambda_D \sim 7 \times 10^{-5}$  m, one finds  $\epsilon_{\lambda_D} \sim \epsilon_c \sim 10^{-3}$ .

Noting the sizes of these small parameters, a common gyrokinetic ordering applicable to the core, and the one used in this work, equates all but  $\epsilon_\tau$  to the normalised gyroradius. This gyrokinetic ordering can be summarised via

$$\epsilon_\delta \sim \epsilon_\omega \sim \epsilon_{\perp} \sim \epsilon_{\lambda_D} \sim \epsilon_c \sim \rho_* \quad (3.9)$$

with  $\epsilon_\tau \sim \rho_*^3$ . The disparate nature of the scales in the system can be exploited to solve equations as an expansion in orders of the smallness parameter  $\rho_*$ .

---

<sup>4</sup>Some works may instead use the tokamak major radius  $R_0$ .

## 3.2 Equations of motion

The analysis of particle motion in turbulent fields is an essential ingredient in the derivation of the gyrokinetic equation, and is typically based on an approach in which one changes phase coordinates to describe the gyrocentre  $\mathbf{R}$ , rather than the particle position  $\mathbf{r}$ . Then by averaging over the gyrophase, the rapid timescale of the gyromotion is removed, and so becomes more adept to describing phenomena with the frequency of the fluctuations,  $\mathcal{O}(\rho_*\Omega_s = v_{\text{th}}/a)$ . Modern derivations of the equations of motion can use a Hamiltonian formulation [72, 73], however in this work we consider the iterative approach.

### 3.2.1 Phase coordinates

#### Constant field summary

The analysis of charged particle motion in constant fields from section 1.2.4 introduced several essential features of single particle motion. It was seen that the dynamics of a charged particle in that simple case consisted of three aspects, namely streaming parallel to the magnetic field, the  $E \times B$  drift, and the gyromotion, which is characterised by the gyrofrequency  $\Omega_s$  and the gyroradius  $\rho_s$ . The rapidity of the rotation relative to the other dynamical processes motivated the idea of the gyrocentre  $\mathbf{R}$ , defined as the instantaneous centre of the gyromotion of a charged particle, which was seen to vary only due to the parallel streaming and  $E \times B$  drifts (equation 1.44). Additionally, it was shown that the magnitude of the velocity associated with the gyromotion  $|\mathbf{u}|$  and the energy  $\varepsilon$  are conserved. These foundational concepts will now be built upon as we consider the turbulent field case.

#### General definition of the gyrocentre

Having established the orderings of the physical quantities in gyrokinetics, we now transform phase coordinates to describe the gyrocentre  $\mathbf{R}$ . The centre of the ring traced out by the gyromotion, which we parameterise using the gyrophase  $\varphi$ , while holding all other quantities including the gyrocentre itself constant *defines* the gyrocentre. Mathematically, this is

$$\mathbf{R} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{r}(t, \mathbf{R}, v_1, v_2, \varphi) d\varphi \quad (3.10)$$

where  $v_1$  and  $v_2$  are as-yet unspecified velocity coordinates. This equation defines both the gyrocentre and the gyrophase, and is an example of a *gyroaverage*, taken here at constant  $\mathbf{R}$ . As a shorthand, we introduce  $\langle \dots \rangle_{\mathbf{R}}$ , such that  $\mathbf{R} = \langle \mathbf{r} \rangle_{\mathbf{R}}$ . The particle position can then be

split into the gyrocentre and the gyroradius,

$$\mathbf{r} = \mathbf{R} + \boldsymbol{\rho}(t, \mathbf{R}, v_1, v_2, \varphi) \quad (3.11)$$

where  $\boldsymbol{\rho} = \mathbf{r} - \langle \mathbf{r} \rangle_{\mathbf{R}}$ . Note that by definition,  $\langle \boldsymbol{\rho} \rangle_{\mathbf{R}} = \mathbf{0}$ .

The definition of the gyrocentre given by equation 3.10 is self-referential, and so cannot be written in closed form. By expanding the equations of motion in orders of  $\rho_*$  however, low-order solutions for its evolution can be found.

### Phase coordinate transformation

Having gyroaveraged the particle position vector, one can also gyroaverage the velocity vector, to separate it into a gyroaveraged part and a non-gyroaveraged part

$$\mathbf{v} = \mathbf{U}(t, \mathbf{R}, v_1, v_2) + \mathbf{u}(t, \mathbf{R}, v_1, v_2, \varphi) \quad (3.12)$$

where  $\mathbf{U} = \langle \mathbf{v} \rangle_{\mathbf{R}}$  and  $\mathbf{u} = \mathbf{v} - \langle \mathbf{v} \rangle_{\mathbf{R}}$ . Note again that  $\langle \mathbf{u} \rangle_{\mathbf{R}} = \mathbf{0}$ .

We therefore transform our system from  $\{t, \mathbf{r}, \mathbf{v}\}$  to the phase coordinates  $\{t, \mathbf{R}, v_1, v_2, \varphi\}$ . As  $\boldsymbol{\rho}$ ,  $\mathbf{U}$  and  $\mathbf{u}$  are as-yet unknown functions of our phase coordinates, then in line with the perturbation theory approach we expand them in orders of  $\rho_*$ . For the gyroradius, we get  $\boldsymbol{\rho} = \boldsymbol{\rho}_0(t, \mathbf{R}, v_1, v_2, \varphi) + \boldsymbol{\rho}_1(t, \mathbf{R}, v_1, v_2, \varphi) + \dots$  where the subscript denotes the size of the term,  $\boldsymbol{\rho}_m \sim \mathcal{O}(\rho_*^m \rho_s)$ . Similarly for  $\mathbf{U}$  and  $\mathbf{u}$ , we have  $\mathbf{U} = \mathbf{U}_0(t, \mathbf{R}, v_1, v_2) + \mathbf{U}_1(t, \mathbf{R}, v_1, v_2) + \dots$  and  $\mathbf{u} = \mathbf{u}_0(t, \mathbf{R}, v_1, v_2, \varphi) + \mathbf{u}_1(t, \mathbf{R}, v_1, v_2, \varphi) + \dots$  where  $\mathbf{U}_m, \mathbf{u}_m \sim \mathcal{O}(\rho_*^m v_{th,s})$ . Our phase coordinate transformations are therefore

$$t = t \quad (3.13)$$

$$\mathbf{r} = \mathbf{R} + \boldsymbol{\rho}_0(t, \mathbf{R}, v_1, v_2, \varphi) + \boldsymbol{\rho}_1(t, \mathbf{R}, v_1, v_2, \varphi) + \dots \quad (3.14)$$

$$\mathbf{v} = \mathbf{U}_0(t, \mathbf{R}, v_1, v_2) + \mathbf{u}_0(t, \mathbf{R}, v_1, v_2, \varphi) + \mathbf{U}_1(t, \mathbf{R}, v_1, v_2) + \mathbf{u}_1(t, \mathbf{R}, v_1, v_2, \varphi) + \dots \quad (3.15)$$

with inverse transformations  $\mathbf{R} = \mathbf{r} - \boldsymbol{\rho}_0(t, \mathbf{r}, \mathbf{v}) - \boldsymbol{\rho}_1(t, \mathbf{r}, \mathbf{v}) - \dots$ ,  $v_1 = v_1(t, \mathbf{r}, \mathbf{v})$ ,  $v_2 = v_2(t, \mathbf{r}, \mathbf{v})$  and  $\varphi = \varphi(t, \mathbf{r}, \mathbf{v})$ . The total time derivative in our phase coordinates is

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \frac{\partial}{\partial \mathbf{R}} + \frac{dv_1}{dt} \frac{\partial}{\partial v_1} + \frac{dv_2}{dt} \frac{\partial}{\partial v_2} + \frac{d\varphi}{dt} \frac{\partial}{\partial \varphi}. \quad (3.16)$$

All other quantities in the system are now ordered. For spatial derivatives<sup>5</sup> acting on a fluctuating quantity  $\delta g$ , then from section 3.1.2 the sizes of the derivatives depend on whether

---

<sup>5</sup>It is shown in appendix B.3 that  $\nabla = \partial/\partial \mathbf{R}$  to lowest order, and thus are typically used somewhat interchangeably.

they are taken parallel or perpendicular to the field  $\nabla = \nabla_{\parallel} + \nabla_{\perp}$ , such that

$$\nabla \delta g = \underbrace{\nabla_{\perp} \delta g}_{\mathcal{O}(\delta g / \rho_s)} + \underbrace{\nabla_{\parallel} \delta g}_{\mathcal{O}(\rho_* \delta g / \rho_s)} \quad (3.17)$$

where  $\nabla_{\parallel} = \hat{\mathbf{b}} \cdot \nabla$  and  $\nabla_{\perp} = \nabla - \nabla_{\parallel} \hat{\mathbf{b}}$ . For equilibrium quantities  $F$ , these all vary on the macroscopic length scale

$$\nabla F = \mathcal{O}(F/a). \quad (3.18)$$

As a consequence of the gyrokinetic orderings we have that the gyromotion is the fastest process in the system, such that

$$\frac{d\varphi}{dt} = \mathcal{O}(\Omega_s) \quad (3.19)$$

and for all processes without gyrophase dependence

$$\frac{d}{dt} = \mathcal{O}(\rho_* \Omega_s). \quad (3.20)$$

Quantities generally contain terms of different orders, and so we can denote these with a subscript. For example,

$$\frac{dg}{dt} = \left( \frac{dg}{dt} \right)_0 + \left( \frac{dg}{dt} \right)_1 + \dots \quad (3.21)$$

where  $(dg/dt)_0 = \mathcal{O}(\Omega_s g)$ ,  $(dg/dt)_1 = \mathcal{O}(\rho_* \Omega_s g)$ . This means then for some gyrophase-dependent quantity  $g(t, \mathbf{R}, v_1, v_2, \varphi)$  we have

$$\frac{dg}{dt} = \frac{\partial g}{\partial \varphi} \left( \frac{d\varphi}{dt} \right)_0 + \mathcal{O}(\rho_* \Omega_s g). \quad (3.22)$$

We now apply these phase coordinate transformations and orderings to the electromagnetic fields. As we are retaining only the equilibrium magnetic field and the fluctuating electric field, the magnetic field expands using equation 3.14 via

$$\mathbf{B}(\mathbf{r}, t) = \underbrace{\mathbf{B}(\mathbf{R}, t)}_{\mathcal{O}(B)} + \underbrace{(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}|_{\mathbf{R}, t}}_{\mathcal{O}(\rho_* B)} + \mathcal{O}(\rho_*^2 B). \quad (3.23)$$

For the electric field, we consider the implications of imposing the gyrokinetic orderings on the fluctuating component of Faraday's law. Expressing the fluctuating electric field in terms of the fluctuating electrostatic potential, where by our assumptions the lowest order fluctuating field is  $\delta \mathbf{E}_1 \sim \mathcal{O}(\rho_* v_{\text{th},s} B)$ ,

$$\delta \mathbf{E}_1 + \delta \mathbf{E}_2 + \dots = -(\nabla_{\perp} + \nabla_{\parallel}) (\delta \phi_1 + \delta \phi_2 + \dots) \quad (3.24)$$

one obtains the lowest order part as

$$\delta \mathbf{E}_1 = -\nabla_{\perp} \delta \phi_1 \quad (3.25)$$

where  $\delta\phi_1 \sim \mathcal{O}(\rho_1^2 v_{\text{th},i} B/a)$ . In our new phase coordinates therefore,

$$\begin{aligned} \delta\mathbf{E}_1(\mathbf{r}, t) &= \underbrace{\delta\mathbf{E}_1(\mathbf{R} + \boldsymbol{\rho}_0(t, \mathbf{R}, v_1, v_2, \varphi), t)}_{\mathcal{O}(\rho_* v_{\text{th},s} B)} + \mathcal{O}(\rho_*^2 v_{\text{th}} B) \\ &= -\nabla_{\perp} \delta\phi_1|_{\mathbf{R} + \boldsymbol{\rho}_0(t, \mathbf{R}, v_1, v_2, \varphi), t}. \end{aligned} \quad (3.26)$$

Here we must retain the lowest-order gyroradius dependence, as  $\boldsymbol{\rho}_0 \cdot \nabla_{\perp} \sim \mathcal{O}(1)$  for fluctuating quantities.

### Equations of motion

Under the above transformations and orderings, the equations of motion for a charged particle

$$\frac{d\mathbf{r}}{dt} = \mathbf{v} \quad (3.27)$$

$$\frac{d\mathbf{v}}{dt} = \frac{q_s}{m_s} (\tilde{\mathbf{E}} + \mathbf{v} \times \tilde{\mathbf{B}}) \quad (3.28)$$

become, to first order

$$\frac{d\mathbf{R}}{dt} = \underbrace{\mathbf{U}_0 + \mathbf{u}_0 - \left(\frac{d\boldsymbol{\rho}_0}{dt}\right)_0}_{\mathcal{O}(v_{\text{th},s})} + \underbrace{\mathbf{U}_1 + \mathbf{u}_1 - \left(\frac{d\boldsymbol{\rho}_0}{dt}\right)_1 - \left(\frac{d\boldsymbol{\rho}_1}{dt}\right)_1}_{\mathcal{O}(\rho_* v_{\text{th},s})} + \mathcal{O}(\rho_*^2 v_{\text{th},s}) \quad (3.29)$$

and

$$\underbrace{\left(\frac{d\mathbf{u}_0}{dt}\right)_0}_{\mathcal{O}(\Omega_s v_{\text{th},s})} + \underbrace{\left(\frac{d\mathbf{U}_0}{dt}\right)_1 + \left(\frac{d\mathbf{u}_0}{dt}\right)_1 + \left(\frac{d\mathbf{u}_1}{dt}\right)_1}_{\mathcal{O}(\rho_* \Omega_s v_{\text{th},s})} = \underbrace{\frac{1}{m_s} \mathbf{F}_0}_{\mathcal{O}(\Omega_s v_{\text{th},s})} + \underbrace{\frac{1}{m_s} \mathbf{F}_1}_{\mathcal{O}(\rho_* \Omega_s v_{\text{th},s})} + \mathcal{O}(\rho_*^2 \Omega_s v_{\text{th},s}) \quad (3.30)$$

respectively, with

$$\mathbf{F}_0 = q_s [(\mathbf{U}_0 + \mathbf{u}_0) \times \mathbf{B}] \quad (3.31)$$

$$\mathbf{F}_1 = q_s [-\nabla_{\perp} \delta\phi_1 + (\mathbf{U}_1 + \mathbf{u}_1) \times \mathbf{B} + (\mathbf{U}_0 + \mathbf{u}_0) \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}]]. \quad (3.32)$$

### 3.2.2 Lowest order solution

#### Force equation

The lowest order part of equation 3.30 is

$$\frac{\partial \mathbf{u}_0}{\partial \varphi} \left(\frac{d\varphi}{dt}\right)_0 = \frac{q_s}{m_s} [(\mathbf{U}_0 + \mathbf{u}_0) \times \mathbf{B}]. \quad (3.33)$$

From our assumption that every process is slower than the gyrofrequency, this includes the evolution of the gyrofrequency, and thus the lowest order part of its evolution should be

independent of gyrophase,  $(\partial/\partial\varphi)[(d\varphi/dt)_0] = 0$ . Therefore by gyroaveraging both sides at constant  $\mathbf{R}$ , we find  $\mathbf{U}_0 \times \mathbf{B} = \mathbf{0}$ , which has the solution

$$\mathbf{U}_0 = U_{0,\parallel}(t, \mathbf{R}, v_1, v_2) \hat{\mathbf{b}}(\mathbf{R}, t) \quad (3.34)$$

where  $\hat{\mathbf{b}}(\mathbf{R}, t) = \mathbf{B}(\mathbf{R}, t)/B(\mathbf{R}, t)$ , for equilibrium field magnitude  $B$ . Equation 3.33 hence becomes

$$\frac{\partial \mathbf{u}_0}{\partial \varphi} \left( \frac{d\varphi}{dt} \right)_0 = \frac{q_s}{m_s} [\mathbf{u}_0 \times \mathbf{B}]. \quad (3.35)$$

To solve this, we introduce a local right-handed orthonormal basis  $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{b}}\}$ , each defined at  $\mathbf{R}$  and  $t$ . We first note that by taking the dot product of both sides with  $\mathbf{u}_0$  that  $|\mathbf{u}_0| = u_0$  is independent of  $\varphi$ . A similar result is obtained from its parallel component,  $u_{0,\parallel} = \mathbf{u}_0 \cdot \hat{\mathbf{b}}$ ,

$$\frac{\partial(\mathbf{u}_0 \cdot \mathbf{u}_0)}{\partial \varphi} = 0, \quad \frac{\partial(\mathbf{u}_0 \cdot \hat{\mathbf{b}})}{\partial \varphi} = 0. \quad (3.36)$$

Because by definition  $\langle \mathbf{u}_0 \rangle_{\mathbf{R}} = \mathbf{0}$ , this implies that  $u_{0,\parallel} = 0$ . We define the unit vector in the direction of  $\mathbf{u}_0$  as  $\hat{\mathbf{e}}_{\perp}$ , such that  $\mathbf{u}_0 = u_0 \hat{\mathbf{e}}_{\perp}$ , where  $\hat{\mathbf{e}}_{\perp} \cdot \hat{\mathbf{b}} = 0$ . Equation 3.35 thus becomes

$$\left( \frac{d\varphi}{dt} \right)_0 \frac{\partial \hat{\mathbf{e}}_{\perp}}{\partial \varphi} = \frac{q_s B}{m_s} \hat{\mathbf{e}}_{\perp} \times \hat{\mathbf{b}} \quad (3.37)$$

which from dimensional arguments, implies

$$\left( \frac{d\varphi}{dt} \right)_0 = \frac{q_s B}{m_s} = \Omega_s \quad (3.38)$$

and so the lowest order variation in the gyrophase is the local gyrofrequency, as expected.

The remaining equation is

$$\frac{\partial \hat{\mathbf{e}}_{\perp}}{\partial \varphi} = \hat{\mathbf{e}}_{\perp} \times \hat{\mathbf{b}} \quad (3.39)$$

which is an equation for simple harmonic motion, an analogous version of which was studied previously (equation 1.35). Without loss of generality, we set the integration constants such that

$$\hat{\mathbf{e}}_{\perp} = (\cos(\varphi) \hat{\mathbf{e}}_1 - \sin(\varphi) \hat{\mathbf{e}}_2) \quad (3.40)$$

and thus

$$\mathbf{u}_0 = u_0 \hat{\mathbf{e}}_{\perp} = u_0 (\cos(\varphi) \hat{\mathbf{e}}_1 - \sin(\varphi) \hat{\mathbf{e}}_2). \quad (3.41)$$

## Velocity equation

Using these results in the lowest order part of equation 3.29, one has

$$\frac{d\mathbf{R}}{dt} = U_{0,\parallel} \hat{\mathbf{b}} + u_0 \hat{\mathbf{e}}_{\perp} - \Omega_s \frac{\partial \rho_0}{\partial \varphi} + \mathcal{O}(\rho_* v_{\text{th},s}). \quad (3.42)$$

Because by our ordering assumptions the gyrocentre velocity should not vary on the timescale of the gyroorbit, one must have

$$u_0 \hat{\mathbf{e}}_{\perp} = \Omega_s \frac{\partial \boldsymbol{\rho}_0}{\partial \varphi} \quad (3.43)$$

which upon integrating yields

$$\begin{aligned} \boldsymbol{\rho}_0 &= \frac{u_0}{\Omega_s} [\sin(\varphi) \hat{\mathbf{e}}_1 + \cos(\varphi) \hat{\mathbf{e}}_2] + \mathbf{C}(t, \mathbf{R}, v_1, v_2) \\ &= \frac{u_0}{\Omega_s} \hat{\mathbf{b}} \times \hat{\mathbf{e}}_{\perp} \end{aligned} \quad (3.44)$$

where  $\mathbf{C} = \mathbf{0}$  in order to satisfy  $\langle \boldsymbol{\rho}_0 \rangle_{\mathbf{R}} = \mathbf{0}$ .

To recap our results thus far, after having solved for the lowest order equations of motion, we have obtained the phase coordinate transformations

$$\mathbf{r} = \mathbf{R} + \frac{u_0}{\Omega_s} \hat{\mathbf{b}} \times \hat{\mathbf{e}}_{\perp} + \mathcal{O}(\rho_* \rho_s) \quad (3.45)$$

$$\mathbf{v} = U_{0,\parallel} \hat{\mathbf{b}} + u_0 \hat{\mathbf{e}}_{\perp} + \mathcal{O}(\rho_* v_{\text{th},s}) \quad (3.46)$$

where  $U_{0,\parallel}$  and  $u_0$  remain undetermined to this order. These variables have been labelled as such for the sake of the generality of the ordering and delineation between gyro-averaged and non-gyro-averaged parts, however we note that these same quantities are often labelled  $v_{\parallel}$  and  $v_{\perp}$  in other works. This is the Catto transformation [74]. The gyrocentre velocity to lowest order is

$$\frac{d\mathbf{R}}{dt} = U_{0,\parallel} \hat{\mathbf{b}} + \mathcal{O}(\rho_* v_{\text{th},s}) \quad (3.47)$$

and the evolution of the gyrophase is

$$\frac{d\varphi}{dt} = \Omega_s + \mathcal{O}(\rho_* \Omega_s). \quad (3.48)$$

We note that these results closely resemble those of the constant field solution given in section 1.2.4 with no equilibrium electric field.

### 3.2.3 First order

We now consider the equations of motion to next order. For the gyrocentre velocity, equation 3.29, including the first order contribution reads

$$\frac{d\mathbf{R}}{dt} = \underbrace{U_{0,\parallel} \hat{\mathbf{b}}}_{\mathcal{O}(v_{\text{th},s})} + \underbrace{\mathbf{U}_1 + \mathbf{u}_1 - \left( \frac{d\boldsymbol{\rho}_0}{dt} \right)_1 - \Omega_s \frac{\partial \boldsymbol{\rho}_1}{\partial \varphi}}_{\mathcal{O}(\rho_* v_{\text{th},s})} + \mathcal{O}(\rho_*^2 v_{\text{th},s}) \quad (3.49)$$

Using equation 3.44, we have

$$\left( \frac{d\boldsymbol{\rho}_0}{dt} \right)_1 = \left( \frac{d}{dt} \left[ \frac{\hat{\mathbf{b}}}{\Omega_s} \right] \right)_1 \times \mathbf{u}_0 + \frac{\hat{\mathbf{b}}}{\Omega_s} \times \left( \frac{d\mathbf{u}_0}{dt} \right)_1 \quad (3.50)$$

and so the equation becomes

$$\frac{d\mathbf{R}}{dt} = \underbrace{U_{0,\parallel}\hat{\mathbf{b}}}_{\mathcal{O}(v_{\text{th},s})} + \underbrace{\mathbf{U}_1 + \mathbf{u}_1 - \left(\frac{d}{dt}\left[\frac{\mathbf{b}}{\Omega_s}\right]\right)_1 \times \mathbf{u}_0 - \frac{\hat{\mathbf{b}}}{\Omega_s} \times \left(\frac{d\mathbf{u}_0}{dt}\right)_1 - \Omega_s \frac{\partial \rho_1}{\partial \varphi}}_{\mathcal{O}(\rho_* v_{\text{th},s})} + \mathcal{O}(\rho_*^2 v_{\text{th},s}). \quad (3.51)$$

Ultimately we are interested in the equations of motion averaged over the gyrophase, as the gyrophase-level variation is unimportant. Taking the gyroaverage, we obtain

$$\left\langle \frac{d\mathbf{R}}{dt} \right\rangle_{\mathbf{R}} = \underbrace{U_{0,\parallel}\hat{\mathbf{b}}}_{\mathcal{O}(v_{\text{th},s})} + \underbrace{\mathbf{U}_1 - \frac{\hat{\mathbf{b}}}{\Omega_s} \times \left\langle \left(\frac{d\mathbf{u}_0}{dt}\right)_1 \right\rangle_{\mathbf{R}}}_{\mathcal{O}(\rho_* v_{\text{th},s})} + \mathcal{O}(\rho_*^2 v_{\text{th},s}). \quad (3.52)$$

To evaluate these first order terms we now consider equation 3.30.

### Force equation

The first order part of equation 3.30 is

$$\left(\frac{d\mathbf{U}_0}{dt}\right)_1 + \left(\frac{d\mathbf{u}_0}{dt}\right)_1 + \frac{\partial \mathbf{u}_1}{\partial \varphi} \Omega_s = \frac{1}{m_s} \mathbf{F}_1. \quad (3.53)$$

Taking the gyroaverage and rearranging, we have

$$\left\langle \left(\frac{d\mathbf{u}_0}{dt}\right)_1 \right\rangle_{\mathbf{R}} = \frac{1}{m_s} \langle \mathbf{F}_1 \rangle_{\mathbf{R}} - \left\langle \left(\frac{d\mathbf{U}_0}{dt}\right)_1 \right\rangle_{\mathbf{R}}. \quad (3.54)$$

Evaluating the gyroaverage of force term given by equation 3.32 yields

$$\frac{1}{m_s} \langle \mathbf{F}_1 \rangle_{\mathbf{R}} = \frac{q_s}{m_s} [-\langle \nabla_{\perp} \delta \phi_1 \rangle_{\mathbf{R}} + \mathbf{U}_1 \times \mathbf{B} + \langle \mathbf{u}_0 \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}] \rangle_{\mathbf{R}}]. \quad (3.55)$$

Here the gyroaverage cannot commute with the gradient of  $\delta \phi_1$ , and so using  $\nabla = \partial / \partial \mathbf{R} + \mathcal{O}(\rho_* k_{\perp})$  from equation B.22 we may write

$$\langle \nabla_{\perp} \delta \phi_1 \rangle_{\mathbf{R}} = \frac{\partial \langle \delta \phi_1 \rangle_{\mathbf{R}}}{\partial \mathbf{R}} + \mathcal{O}(\rho_* k_{\perp}). \quad (3.56)$$

In appendix B.2, it is shown that the third term of equation 3.55 averages to

$$\langle \mathbf{u}_0 \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}] \rangle_{\mathbf{R}} = -\frac{u_0^2}{2\Omega_s} \nabla B. \quad (3.57)$$

For the second term in equation 3.54, one obtains

$$\left\langle \left(\frac{d\mathbf{U}_0}{dt}\right)_1 \right\rangle_{\mathbf{R}} = \left\langle \left(\frac{dU_{0,\parallel}}{dt}\right)_1 \right\rangle_{\mathbf{R}} \hat{\mathbf{b}} + U_{0,\parallel}^2 (\hat{\mathbf{b}} \cdot \nabla) \hat{\mathbf{b}}. \quad (3.58)$$

Taking the cross product of equation 3.54 with  $-\hat{\mathbf{b}}/\Omega_s$ , adding  $\mathbf{U}_1$  to both sides and using equations 3.55, 3.56, 3.57 and 3.58, we obtain

$$\mathbf{U}_1 - \frac{1}{\Omega_s} \hat{\mathbf{b}} \times \left\langle \left(\frac{d\mathbf{u}_0}{dt}\right)_1 \right\rangle_{\mathbf{R}} = U_{1,\parallel} \hat{\mathbf{b}} + \frac{\hat{\mathbf{b}}}{B} \times \frac{\partial \langle \delta \phi_1 \rangle_{\mathbf{R}}}{\partial \mathbf{R}} + \frac{U_{0,\parallel}^2}{\Omega_s} \hat{\mathbf{b}} \times [(\hat{\mathbf{b}} \cdot \nabla) \hat{\mathbf{b}}] + \frac{u_0^2}{2\Omega_s B} \hat{\mathbf{b}} \times \nabla B \quad (3.59)$$

which when substituted into equation 3.52, gives

$$\begin{aligned} \left\langle \frac{d\mathbf{R}}{dt} \right\rangle_{\mathbf{R}} &= \underbrace{U_{0,\parallel} \hat{\mathbf{b}}}_{\mathcal{O}(v_{\text{th},s})} + U_{1,\parallel} \hat{\mathbf{b}} + \underbrace{\frac{\hat{\mathbf{b}}}{B} \times \frac{\partial \langle \delta \phi_1 \rangle_{\mathbf{R}}}{\partial \mathbf{R}} + \frac{U_{0,\parallel}^2}{\Omega_s} \hat{\mathbf{b}} \times \left[ \left( \hat{\mathbf{b}} \cdot \nabla \right) \hat{\mathbf{b}} \right] + \frac{u_0^2}{2\Omega_s B} \hat{\mathbf{b}} \times \nabla B}_{\mathcal{O}(\rho_* v_{\text{th},s})} + \mathcal{O}(\rho_*^2 v_{\text{th},s}) \\ &= U_{0,\parallel} \hat{\mathbf{b}} + U_{1,\parallel} \hat{\mathbf{b}} + \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} + \mathbf{v}_D + \mathcal{O}(\rho_*^2 v_{\text{th},s}) \end{aligned} \quad (3.60)$$

which is our desired result, describing the gyroaveraged evolution of the gyrocentre  $\mathbf{R}$  for a charged particle in an equilibrium magnetic field and a fluctuating electric field. Here  $U_{0,\parallel}$  and  $U_{1,\parallel}$  are the lowest and first order parallel velocities<sup>6</sup>,  $\langle \delta \mathbf{v}_E \rangle_{\mathbf{R}}$  is the gyroaveraged turbulent  $E \times B$  velocity,

$$\langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} = \frac{\hat{\mathbf{b}}}{B} \times \frac{\partial \langle \delta \phi_1 \rangle_{\mathbf{R}}}{\partial \mathbf{R}} \quad (3.61)$$

and  $\mathbf{v}_D$  describes the equilibrium magnetic drifts,

$$\mathbf{v}_D = \frac{U_{0,\parallel}^2}{\Omega_s} \hat{\mathbf{b}} \times \left[ \left( \hat{\mathbf{b}} \cdot \nabla \right) \hat{\mathbf{b}} \right] + \frac{u_0^2}{2\Omega_s B} \hat{\mathbf{b}} \times \nabla B \quad (3.62)$$

which are the curvature drift and the grad-B drift respectively.

### 3.2.4 Velocity coordinates

Equation 3.60 has been derived without specifying the velocity coordinates  $v_1$  and  $v_2$ . This freedom is the reason why not all formulations of the gyrokinetic equation have the same velocity coordinates, for which different choices can be made for different purposes.

‘Natural’ choices for velocity coordinates may be the lowest-order parallel velocity  $U_{0,\parallel}$  and gyromotion speed  $u_0$ , due to their appearance in equation 3.60. Their gyroaveraged evolution equations can be obtained from equation 3.53, shown in appendix B.4 to be

$$\left\langle \frac{dU_{0,\parallel}}{dt} \right\rangle_{\mathbf{R}} = -\frac{u_0^2}{2B} \hat{\mathbf{b}} \cdot \nabla B + \mathcal{O}(\rho_*^2 \Omega_s v_{\text{th},s}) \quad (3.63)$$

$$\left\langle \frac{du_0}{dt} \right\rangle_{\mathbf{R}} = \frac{u_0}{2B} \mathbf{U}_0 \cdot \nabla B + \mathcal{O}(\rho_*^2 \Omega_s v_{\text{th},s}) \quad (3.64)$$

however other velocity coordinates may also be constructed.

### Magnetic moment

Consider the total derivative of the equilibrium magnetic field,

$$\frac{dB}{dt} = \mathbf{U}_0 \cdot \nabla B + \mathcal{O}(\rho_*^2 \Omega_s B). \quad (3.65)$$

---

<sup>6</sup>The first order correction to the parallel velocity  $U_{1,\parallel}$  will be seen not to appear in the gyrokinetic equation, however it is included here for completeness.

We see that this has a similar form to the gyroaveraged evolution of  $u_0$ , equation 3.64. Now consider a function  $\mu = \mu(u_0, B)$ , for which the total time derivative is

$$\frac{d\mu}{dt} = \frac{\partial\mu}{\partial u_0} \frac{du_0}{dt} + \frac{\partial\mu}{\partial B} \frac{dB}{dt}. \quad (3.66)$$

Gyroaveraging both sides, one obtains

$$\begin{aligned} \left\langle \frac{d\mu}{dt} \right\rangle_{\mathbf{R}} &= \frac{\partial\mu}{\partial u_0} \left\langle \frac{du_0}{dt} \right\rangle_{\mathbf{R}} + \frac{\partial\mu}{\partial B} \frac{dB}{dt} \\ &= \left( \frac{\partial\mu}{\partial u_0} \frac{u_0}{2B} + \frac{\partial\mu}{\partial B} \right) \mathbf{U}_0 \cdot \nabla B + \mathcal{O}(\rho_*^2 \Omega_s \mu). \end{aligned} \quad (3.67)$$

The first term of this equation vanishes for  $\mu = C_\mu u_0^2/B$ , where  $C_\mu$  is a constant. Taking  $C_\mu = m_s/2$  defines the ‘magnetic moment’ of the particle<sup>7</sup>,  $\mu = m_s u_0^2/(2B)$ . This quantity is therefore conserved up to second order,

$$\left\langle \frac{d\mu}{dt} \right\rangle_{\mathbf{R}} = \mathcal{O}(\rho_*^2 \Omega_s \mu) \quad (3.68)$$

and so provides a natural choice of velocity coordinate.

## Kinetic energy

Let us now consider conservation of energy. The evolution of the energy is most readily found by first returning to the original coordinate system,  $\{t, \mathbf{r}, \mathbf{v}\}$ . Taking the dot product of the Lorentz force with  $\mathbf{v}$  and re-arranging, one obtains

$$\frac{d}{dt} \left( \frac{1}{2} m_s |\mathbf{v}|^2 \right) = q_s \mathbf{v} \cdot \delta \mathbf{E} \quad (3.69)$$

which is true to all orders. Returning to our phase coordinates, the kinetic energy to lowest order is

$$\varepsilon = \frac{1}{2} m |\mathbf{v}|^2 = \frac{1}{2} m_s \left( U_{0,\parallel}^2 + u_0^2 \right) + \mathcal{O}(\rho_* T_s) \quad (3.70)$$

and the right hand side of 3.69 to second order is simply  $q_s \mathbf{v} \cdot \delta \mathbf{E} = -q_s \mathbf{u}_0 \cdot \nabla_\perp \delta \phi_1 + \mathcal{O}(\rho_*^2 \Omega_s T)$ .

Therefore to lowest order, energy conservation is given by

$$\frac{d\varepsilon}{dt} = -q_s \mathbf{u}_0 \cdot \nabla_\perp \delta \phi_1 + \mathcal{O}(\rho_*^2 \Omega_s T_s) \quad (3.71)$$

which upon gyroaveraging yields

$$\left\langle \frac{d\varepsilon}{dt} \right\rangle_{\mathbf{R}} = \mathcal{O}(\rho_*^2 \Omega_s T_s) \quad (3.72)$$

where we have used equation B.28. Similarly to the magnetic moment, we have found the lowest order kinetic energy is conserved to second order.

<sup>7</sup>Note some authors define the magnetic moment without the factor of mass,  $\mu = u_0^2/(2B)$ .

If one were to make the choice of using the magnetic moment and the kinetic energy for their velocity coordinates, their conservation yields a particularly simple version of the gyrokinetic equation, as will be seen. However, their use is only particularly suited to the collisionless case, such as in the code GYRO [75]. Explicitly resolving the effects of the collision operator is generally better suited to other coordinates. For example, the codes GKW [35] and GENE [76] use the parallel velocity and the magnetic moment, whereas CGYRO [38], a code purpose-built for dealing with collisions, uses the energy and the ‘pitch angle’, the ratio of the parallel velocity to the speed.

### General coordinates

For a general velocity coordinate  $v'$ , one can consider its evolution equation using

$$\frac{dv'}{dt} = \frac{\partial v'}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \frac{\partial v'}{\partial \mathbf{R}} + \frac{dU_{0,\parallel}}{dt} \frac{\partial v'}{\partial U_{0,\parallel}} + \frac{du_0}{dt} \frac{\partial v'}{\partial u_0} + \frac{d\varphi}{dt} \frac{\partial v'}{\partial \varphi}. \quad (3.73)$$

To be slowly-varying at lowest order, one requires  $\partial v'/\partial \varphi = 0$ . Assuming time and position dependence only on equilibrium quantities, this gyroaverages to

$$\left\langle \frac{dv'}{dt} \right\rangle_{\mathbf{R}} = U_{0,\parallel} \left( \frac{\partial v'}{\partial B} + \frac{u_0}{2B} \frac{\partial v'}{\partial u_0} - \frac{u_0^2}{2B U_{0,\parallel}} \frac{\partial v'}{\partial U_{0,\parallel}} \right) \hat{\mathbf{b}} \cdot \nabla B + \mathcal{O}(\rho_*^2 \Omega_s v'). \quad (3.74)$$

which can be used to describe the gyroaveraged evolution of any velocity coordinate to first order.

### 3.2.5 Single particle motion summary

By applying the gyrokinetic ordering assumptions to the Lorentz force equation, the equations of motion for a charged particle in the electromagnetic field can be solved via the transformations

$$\mathbf{r} = \mathbf{R} + \frac{1}{\Omega_s} \hat{\mathbf{b}} \times \mathbf{u}_0 + \mathcal{O}(\rho_* \rho_s) \quad (3.75)$$

$$\mathbf{v} = \mathbf{U}_0 + \mathbf{u}_0 + \mathcal{O}(\rho_* v_{\text{th},s}) \quad (3.76)$$

where  $\mathbf{U}_0$  and  $\mathbf{u}_0$  are given by equations 3.34 and 3.41. The gyro-averaged evolution equations of the gyrocentre and the gyrophase are

$$\left\langle \frac{d\mathbf{R}}{dt} \right\rangle_{\mathbf{R}} = U_{0,\parallel} \hat{\mathbf{b}} + U_{1,\parallel} \hat{\mathbf{b}} + \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} + \mathbf{v}_D + \mathcal{O}(\rho_*^2 v_{\text{th},s}) \quad (3.77)$$

$$\left\langle \frac{d\varphi}{dt} \right\rangle_{\mathbf{R}} = \Omega_s + \mathcal{O}(\rho_* \Omega_s) \quad (3.78)$$

where  $U_{1,\parallel}$  is the first order parallel velocity,  $\langle \delta \mathbf{v}_E \rangle_{\mathbf{R}}$  is the turbulent  $E \times B$  velocity (equation 3.61) and  $\mathbf{v}_D$  comprises the equilibrium magnetic drifts (equation 3.62). The equations of motion for the lowest order velocity functions are

$$\left\langle \frac{dU_{0,\parallel}}{dt} \right\rangle_{\mathbf{R}} = -\frac{u_0^2}{2B} \hat{\mathbf{b}} \cdot \nabla B + \mathcal{O}(\rho_*^2 \Omega_s v_{\text{th},s}) \quad (3.79)$$

$$\left\langle \frac{du_0}{dt} \right\rangle_{\mathbf{R}} = \frac{u_0}{2B} \mathbf{U}_0 \cdot \nabla B + \mathcal{O}(\rho_*^2 \Omega_s v_{\text{th},s}) \quad (3.80)$$

however velocity coordinates can be found which are conserved up to second order. These are the lowest order kinetic energy  $\varepsilon = \frac{1}{2} m_s (U_{0,\parallel}^2 + u_0^2)$  and the magnetic moment  $\mu = m_s u_0^2 / (2B)$ , which satisfy

$$\left\langle \frac{d\varepsilon}{dt} \right\rangle_{\mathbf{R}} = \mathcal{O}(\rho_*^2 \Omega_s T_s) \quad (3.81)$$

$$\left\langle \frac{d\mu}{dt} \right\rangle_{\mathbf{R}} = \mathcal{O}(\rho_*^2 \Omega_s T_s / B). \quad (3.82)$$

More generally, the gyroaveraged evolution equation for a slowly-varying velocity coordinate  $v' = v'(t, \mathbf{R}, U_{0,\parallel}, u_0)$  can be found from

$$\left\langle \frac{dv'}{dt} \right\rangle_{\mathbf{R}} = U_{0,\parallel} \left( \frac{\partial v'}{\partial B} + \frac{u_0}{2B} \frac{\partial v'}{\partial u_0} - \frac{u_0^2}{2B} \frac{1}{U_{0,\parallel}} \frac{\partial v'}{\partial U_{0,\parallel}} \right) \hat{\mathbf{b}} \cdot \nabla B + \mathcal{O}(\rho_*^2 \Omega_s v'). \quad (3.83)$$

We shall now use these results in deriving the gyrokinetic equation [34, 77].

### 3.3 The gyrokinetic equation

For a pedagogical introduction to gyrokinetics, readers are recommended [78] and [79]<sup>8</sup>.

A statistical description of a plasma species  $s$  uses the distribution function,  $f_s$ , which has the property that  $f_s d^3 \mathbf{r} d^3 \mathbf{v}$  is the probability of finding a particle in that volume of phase space. This is a seven-dimensional function, evolving in time, three spatial coordinates and three velocity coordinates. Using the chain rule, its total time derivative in arbitrary phase coordinates  $z^i$  is

$$\frac{df_s}{dt} = \frac{\partial f_s}{\partial t} + z^i \frac{\partial f_s}{\partial z^i} = C_s \quad (3.84)$$

where we have used the Einstein summation convention,  $C_s$  is the collision operator [80–82], and we have used dot notation as a shorthand to denote total time derivatives of  $z^i$ . Having derived the equations of motion in the turbulent system in the previous section, we shall use the phase coordinates  $\{t, \mathbf{R}, v_1, v_2, \varphi\}$ . Explicitly therefore, the kinetic equation is

$$\frac{\partial f_s}{\partial t} + \frac{d\mathbf{R}}{dt} \cdot \frac{\partial f_s}{\partial \mathbf{R}} + \frac{dv_1}{dt} \frac{\partial f_s}{\partial v_1} + \frac{dv_2}{dt} \frac{\partial f_s}{\partial v_2} + \frac{d\varphi}{dt} \frac{\partial f_s}{\partial \varphi} = C_s. \quad (3.85)$$

---

<sup>8</sup>Note that in [79] a uniform background distribution is considered and thus the result is not applicable to tokamaks, however the foundational concepts of gyrokinetics are well-explained.

As the distribution function is part of a turbulent system, it too has an ensemble averaged and fluctuating part,  $f_s = F_s + \delta f_s$ , where  $F_s = \langle f_s \rangle_{\text{Ens}}$  is labelled as such by convention. The distribution function is also subject to ordering, such that we have  $f_s = F_{s,0} + F_{s,1} + \dots + \delta f_{s,1} + \delta f_{s,2} + \dots$  where the fluctuating component is ordered small relative to the equilibrium,  $\delta f_{s,1} \sim \mathcal{O}(\rho_* F_{s,0})$ . The derivatives of the phase coordinates also have equilibrium and fluctuating parts,  $\dot{z}^i = \langle \dot{z}^i \rangle_{\text{Ens}} + \delta \dot{z}^i$ , and are subject to ordering. The majority of this analysis has been carried out in the previous section. The collisions are ordered  $C_s = \mathcal{O}(\rho_* \Omega_s)$ . Before applying the orderings to the kinetic equation, we separate it into its equilibrium and fluctuating parts. Taking the ensemble average of equation 3.84, one has

$$\frac{\partial F_s}{\partial t} + \langle \dot{z}^i \rangle_{\text{Ens}} \frac{\partial F_s}{\partial z^i} + \left\langle \frac{\partial \delta f_s}{\partial z^i} \delta \dot{z}^i \right\rangle_{\text{Ens}} = \langle C_s \rangle_{\text{Ens}} \quad (3.86)$$

and thus the kinetic equation for the fluctuations is

$$\frac{\partial \delta f_s}{\partial t} + \langle \dot{z}^i \rangle_{\text{Ens}} \frac{\partial \delta f_s}{\partial z^i} + \delta \dot{z}^i \frac{\partial F_s}{\partial z^i} + \delta \left[ \frac{\partial \delta f_s}{\partial z^i} \delta \dot{z}^i \right] = \delta C_s \quad (3.87)$$

where we have defined  $\delta \left[ \frac{\partial \delta f_s}{\partial z^i} \delta \dot{z}^i \right] = \frac{\partial \delta f_s}{\partial z^i} \delta \dot{z}^i - \left\langle \frac{\partial \delta f_s}{\partial z^i} \delta \dot{z}^i \right\rangle_{\text{Ens}}$ . Like with the equations of motion in the previous section, we will now solve the ensemble averaged and fluctuating kinetic equations order by order in  $\rho_*$ .

### 3.3.1 Zeroth order: gyrophase-independence of $F_{s,0}$

The zeroth order part of equation 3.86, namely the terms of size  $\mathcal{O}(\Omega_s F_{s,0})$ , is simply

$$\Omega_s \frac{\partial F_{s,0}}{\partial \varphi} = 0 \quad (3.88)$$

and so one has  $F_{s0} = F_{s0}(t, \mathbf{R}, v_1, v_2)$ , indicating the lowest order ensemble averaged part of the distribution function does not vary on the scale of the gyrofrequency, as expected. There is no zeroth order part of equation 3.87, as all terms are size  $\mathcal{O}(\rho_* F_{s,0} \Omega_s)$  or smaller.

### 3.3.2 First order: Maxwellian background and the splitting of $\delta f_{s,1}$

In the following we choose to specify the velocity coordinates  $v_1 = \varepsilon$  and  $v_2 = \mu$ , due to the integral role of the particle energy in the following and the simplicity achieved by having both of their gyroaveraged evolution equations be conserved up to second order<sup>9</sup>.

<sup>9</sup>The gyrokinetic system can be derived for general velocity coordinates  $v_1$  and  $v_2$ , however the manipulation of partial derivatives is significantly more intricate. Here the result is derived for energy and magnetic moment, before transforming to arbitrary velocity coordinates later in appendix B.6.

The first order equilibrium equation is

$$\left\langle \frac{d\mathbf{R}}{dt} \right\rangle_{\text{Ens}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} + \left\langle \frac{d\varepsilon}{dt} \right\rangle_{\text{Ens}} \frac{\partial F_{s,0}}{\partial \varepsilon} + \left\langle \frac{d\mu}{dt} \right\rangle_{\text{Ens}} \frac{\partial F_{s,0}}{\partial \mu} + \left\langle \frac{d\varphi}{dt} \right\rangle_{\text{Ens}} \frac{\partial F_{s,1}}{\partial \varphi} = \langle C_s \rangle_{\text{Ens}} \quad (3.89)$$

which using our results from section 3.2.5 becomes

$$U_{0,\parallel} \hat{\mathbf{b}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} + \left\langle \frac{d\mu}{dt} \right\rangle_{\text{Ens}} \frac{\partial F_{s,0}}{\partial \mu} + \Omega_s \frac{\partial F_{s,1}}{\partial \varphi} = \langle C_s \rangle_{\text{Ens}}. \quad (3.90)$$

Gyroaveraging the equation, we find

$$U_{0,\parallel} \hat{\mathbf{b}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} = \langle C_s \rangle_{\text{Ens},\mathbf{R}}. \quad (3.91)$$

It can be shown [31, 47] that this equation is satisfied by a local Maxwellian equilibrium that does not vary in the direction parallel to the equilibrium magnetic field, hence

$$F_{s,0}(\mathbf{R}, t, \varepsilon) = n_{s,0}(\mathbf{R}, t) \left( \frac{m_s}{2\pi T_{s,0}(\mathbf{R}, t)} \right)^{3/2} e^{-\varepsilon/T_{s,0}(\mathbf{R}, t)} \quad (3.92)$$

with

$$\hat{\mathbf{b}} \cdot \frac{\partial n_{s,0}}{\partial \mathbf{R}} = \hat{\mathbf{b}} \cdot \frac{\partial T_{s,0}}{\partial \mathbf{R}} = 0. \quad (3.93)$$

Therefore  $\partial F_{s,0}/\partial \mu = 0$ . Now taking 3.91 away from 3.90, we have

$$\Omega_s \frac{\partial F_{s,1}}{\partial \varphi} = 0 \quad (3.94)$$

and thus  $F_{s,1}$  is independent of gyrophase,  $F_{s,1} = F_{s,1}(t, \mathbf{R}, \varepsilon, \mu)$ .

### First order fluctuating equation

The first order part of equation 3.87 is

$$\begin{aligned} \Omega_s \frac{\partial \delta f_{s,1}}{\partial \varphi} \Big|_{t,\mathbf{R},\varepsilon,\mu} &= -\delta \dot{\varepsilon} \frac{\partial F_{s,0}}{\partial \varepsilon} \Big|_{t,\mathbf{R},\mu,\varphi} \\ &= -\frac{q_s}{T_{s,0}} F_{s,0} (\mathbf{u}_0 \cdot \nabla_{\perp} \delta \phi_1). \end{aligned} \quad (3.95)$$

This can be integrated to obtain

$$\begin{aligned} \delta f_{s,1} &= -\frac{q_s}{\Omega_s T_{s,0}} F_{s,0} \int (\mathbf{u}_0 \cdot \nabla_{\perp} \delta \phi_1) d\varphi + h_s(t, \mathbf{R}, \varepsilon, \mu) \\ &= -\frac{q_s}{T_{s,0}} F_{s,0} \int \frac{\partial \delta \phi_1}{\partial \varphi} d\varphi + h_s(t, \mathbf{R}, \varepsilon, \mu) \\ &= -\frac{q_s}{T_{s,0}} F_{s,0} \delta \phi_1 + h_s(t, \mathbf{R}, \varepsilon, \mu) \end{aligned} \quad (3.96)$$

where we have used B.26. Thus we see the first order part of the fluctuating distribution  $\delta f_{s,1}$  is made up of two parts: the adiabatic response, proportional to the fluctuating potential  $\delta \phi_1$ , and the non-adiabatic response  $h_s$ , which is independent of gyrophase.

### 3.3.3 Second order: the gyrokinetic equation

The second order equilibrium equation gives the neoclassical equilibrium equation [83], which describes the evolution of  $F_{s,1}$ . The first-order neoclassical correction to the background Maxwellian does not appear in the gyrokinetic equation, and so  $F_{s,1}$  is not needed in this derivation and is thus beyond the scope of this work.

The gyrokinetic equation is obtained by taking the gyroaverage of the second order fluctuating kinetic equation, which is done in detail in appendix B.5. This describes the evolution of the non-adiabatic part of the fluctuating distribution function  $h_s$ , and is given in the  $\{t, \mathbf{R}, \varepsilon, \mu, \varphi\}$  system by

$$\frac{\partial h_s}{\partial t} + [U_{0,\parallel} \hat{\mathbf{b}} + \mathbf{v}_D] \cdot \frac{\partial h_s}{\partial \mathbf{R}} + \delta \left( \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} \right) = \frac{q_s F_{s,0}}{T_{s,0}} \frac{\partial \langle \delta \phi_1 \rangle_{\mathbf{R}}}{\partial t} - \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} + \langle (\delta C_s)_2 \rangle_{\mathbf{R}} \quad (3.97)$$

which is in agreement with [31, 71, 84] in the limit of no equilibrium flows and no electromagnetic fluctuations. The first term is simply the change in the non-adiabatic distribution with time. The second term encapsulates the effect of the parallel streaming and magnetic drifts on the distribution. The third term is the nonlinear term, which describes the effect of the fluctuating electrostatic potential on  $h_s$ . On the right hand side, one has the change in the gyroaveraged potential with time, the driving term of the gradients for the background Maxwellian distribution and the collision operator.

The gyrokinetic equation shown here is for the  $\{t, \mathbf{R}, \varepsilon, \mu, \varphi\}$  system, however as noted previously, one can choose to express it using different velocity coordinates. The general form of the gyrokinetic equation in a  $\{t, \mathbf{R}, v_1, v_2, \varphi\}$  system is given in appendix B.6. As this work focuses less on explicitly solving the equation, but the consequence of the resulting fluxes, we are content to proceed with  $\varepsilon$  and  $\mu$  description.

The linear gyrokinetic equation can be obtained by a subsequent expansion in perturbation size, which amounts to the nonlinear term being negligible. This is

$$\frac{\partial h_s}{\partial t} + [U_{0,\parallel} \hat{\mathbf{b}} + \mathbf{v}_D] \cdot \frac{\partial h_s}{\partial \mathbf{R}} = \frac{q_s F_{s,0}}{T_{s,0}} \frac{\partial \langle \delta \phi_1 \rangle_{\mathbf{R}}}{\partial t} - \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} + \langle (\delta C)_2 \rangle_{\mathbf{R}} \quad (3.98)$$

which can be used to study the linear stability of plasmas.

### 3.3.4 Maxwell's equations

Having derived the gyrokinetic equation, to complete the system we require the gyrokinetic version of Maxwell's equations. The assumption of  $\delta \mathbf{B} = \mathbf{0}$  renders the fluctuating compo-

nents of Gauss's law for magnetism and Ampère's law trivial<sup>10</sup>, and their ensemble averaged components are used in the following section on tokamak geometry. The fluctuating component of Faraday's law was solved via equation 3.25, and the ensemble averaged part is trivial for  $\mathbf{E} = \mathbf{0}$ . We therefore consider Gauss's law.

### Equilibrium component of Gauss's law

The lowest order equilibrium component of Gauss's law (equation 3.6) is<sup>11</sup>

$$\rho_0 = 0. \quad (3.99)$$

Using the result of equation 3.92, this becomes

$$\sum_s q_s \int F_{s,0} d^3\mathbf{v} = \sum_s q_s n_{s,0} = 0 \quad (3.100)$$

which states that the equilibrium charge density of the plasma is zero, as expected from quasineutrality. Taking the electron contribution to the right-hand side, we have the condition for specifying bulk neutrality in a plasma,

$$\sum_i Z_i e n_{i,0} = e n_{e,0} \implies \sum_i Z_i \left( \frac{n_{i,0}}{n_{e,0}} \right) = 1. \quad (3.101)$$

where  $i$  labels the ion species.

### Fluctuating Gauss's law

For the lowest order part of the fluctuating component of Gauss' law, equation 3.7, we have

$$-\epsilon_0 \nabla_{\perp}^2 \delta\phi_1 = \delta\rho_1. \quad (3.102)$$

Comparing the sizes of these terms, we note that the right-hand side is  $\delta\rho_1 \sim \mathcal{O}(\rho_* e n_0)$ . The left-hand side is of order

$$-\epsilon_0 \nabla_{\perp}^2 \delta\phi \sim \mathcal{O}(\rho_* \epsilon_{\lambda_D} e n_0) \sim \mathcal{O}(\rho_*^2 e n_0) \quad (3.103)$$

and thus is next order. The fluctuating component of Gauss's law is therefore

$$\sum_s q_s \delta n_{s,1} = 0. \quad (3.104)$$

<sup>10</sup>In the case of  $\delta\mathbf{B} \neq \mathbf{0}$  the scalar fields  $\delta A_{\parallel}$  and  $\delta B_{\parallel}$  are described by the parallel and perpendicular components of the fluctuating Ampère's law respectively.

<sup>11</sup>This is true even in the case of  $\mathbf{E} \neq \mathbf{0}$ , due to the ordering of the Debye length.

Using our result for the first order fluctuating part of the distribution function, equation 3.96, we have

$$\sum_s q_s \int \left( -\frac{q_s}{T_{s,0}} F_{s,0} \delta\phi_1 + h_s \right) d^3\mathbf{v} = 0. \quad (3.105)$$

Note that the fluctuating Maxwell's equations are calculated at the particle position  $\mathbf{r}$ , not the gyrocentre  $\mathbf{R}$ , and thus the velocity integrations must be taken over constant  $\mathbf{r}$ . Carrying out the integration of  $F_{s,0}$ , and gyro-averaging  $h_s$  at constant  $\mathbf{r}$ , the fluctuating quasineutrality condition is

$$\sum_s \left[ -\frac{q_s^2 n_{s,0}}{T_{s,0}} \delta\phi_1 + q_s \int \langle h_s \rangle_{\mathbf{r}} d^3\mathbf{v} \right] = 0. \quad (3.106)$$

For an electrostatic plasma, the gyrokinetic system consists of equation 3.106 and the set of equations 3.97 for each species  $s$ .

### 3.4 Tokamak geometry

In this section, the coordinate system used for simulations of tokamak turbulence is derived [11, 85]. The gyrokinetic system in section 3.3 was derived for an arbitrary spatial coordinate system, but the so-called ‘field-aligned’ system [86] allows both an efficient use of the simulation volume and a natural treatment of the disparate dynamics parallel and perpendicular to the equilibrium magnetic field. Different codes can employ different coordinates and conventions within the field-aligned framework, and this section seeks to unify some of these concepts.

Field-aligned coordinates are generally non-orthogonal, and so the following analysis uses the covariant and contravariant representations of vectors. An introduction to these concepts is provided in appendix C. For general coordinates  $\xi^i$ , we use  $\nabla\xi^i$  for the contravariant basis vectors and  $\mathbf{a}_i = \partial\mathbf{r}/\partial\xi^i$  for the covariant basis vectors, for position vector  $\mathbf{r}$ . The contravariant components of a general vector  $\mathbf{V}$  are therefore  $V^i = \mathbf{V} \cdot \nabla\xi^i$ , with covariant components  $V_i = \mathbf{V} \cdot \mathbf{a}_i$ , which can be related through the components of the metric, for example  $V_i = g_{ij} V^j$ , where we use the Einstein summation convention.

#### 3.4.1 Divergenceless condition

The derivation of the field-aligned system starts from Gauss's Law for magnetism for the equilibrium magnetic field

$$\nabla \cdot \mathbf{B} = 0. \quad (3.107)$$

This allows the magnetic field to be written as the cross product of two gradients

$$\mathbf{B} = \nabla F \times \nabla G \quad (3.108)$$

from equation C.15. This is the Clebsch form [85]. Here  $F$  and  $G$  are two scalar potentials, the gradients of which are always perpendicular to the magnetic field,  $\mathbf{B} \cdot \nabla F = \mathbf{B} \cdot \nabla G = 0$ . Magnetic field lines therefore lie on surfaces of constant  $F$  and surfaces of constant  $G$ , with specific field lines traced out by curves of constant both  $F$  and  $G$ .

These potential functions are as-yet unknown, however if one were to construct a coordinate system with  $F$  and  $G$  as two of the coordinates, the covariant basis vector of the third coordinate  $H$  would be parallel to the magnetic field everywhere, via

$$\mathbf{B} = \nabla F \times \nabla G = \frac{1}{\mathcal{J}_{FGH}} \mathbf{a}_H \quad (3.109)$$

using equation C.30, where  $\mathcal{J}_{FGH}$  is the Jacobian of the coordinate system. This prescribes the parallel direction for a general equilibrium magnetic field configuration to a single coordinate, allowing efficient treatment of the disparate dynamics of plasma turbulence parallel and perpendicular to the field. This also allows when considering the local limit the use of the most computationally-efficient simulation domain for plasma turbulence, which is the ‘local flux tube’. We now solve for the functions  $F$  and  $G$ .

### 3.4.2 Axisymmetric flux surfaces

Coordinate systems that include the toroidal angle are particularly useful in describing tokamaks due to their axisymmetry. We start in a cylindrical coordinate system  $\{R, \Phi, Z\}$ , for which the basis vectors and metric components are given in appendix D.2. Mathematically, axisymmetry means that physical equilibrium quantities, such as the magnetic field components, are independent of toroidal angle  $\partial/\partial\Phi|_{R,Z} = 0$ .

The consequences of axisymmetry on the equilibrium magnetic field are detailed in appendix D.2.1. It is shown that axisymmetry constrains the toroidal angular dependence of  $F$  and  $G$ . One of the potentials becomes independent of  $\Phi$ , which in anticipation of physical interpretation we relabel  $F \rightarrow \varrho$ , such that

$$F = \varrho(R, Z) \quad (3.110)$$

and the second potential has a linear dependence on toroidal angle,

$$G = \bar{c}_1(\varrho) \Phi + \tilde{G}(R, Z). \quad (3.111)$$

The function  $F$  has been relabelled to  $\varrho$  as these surfaces correspond to the *flux surfaces* discussed in section 1.4. This  $\varrho$  forms a label for the flux-surfaces of the tokamak, providing a generalised radial coordinate. Note that any monotonically increasing flux function can hypothetically be used as a radial coordinate, as this uniquely labels each surface. In this work however for generality we will use  $\varrho$ , labelling distance, from  $\varrho = 0$  at the magnetic axis to  $\varrho = a$  at the minor radius at the outboard midplane. The value of the surface potential  $G$  for a given  $\varrho$  therefore labels the magnetic field lines that lie on that surface.

### 3.4.3 Toroidal coordinates

The existence of concentric flux surfaces encourages transformation to toroidal coordinates. These coordinates are  $\{\varrho, \vartheta, \zeta\}$ , shown in figure 3.1, where  $\varrho = \varrho(R, Z)$  is the generalised radial coordinate which labels the flux surfaces,  $\vartheta = \vartheta(R, Z)$  is a generalised poloidal angle, oriented anticlockwise from the outboard midplane, and  $\zeta = -\Phi$  is the toroidal angle, where the minus sign has been introduced to keep the coordinate system right-handed. The specific transformations for  $\varrho$  and  $\vartheta$  are for now left unspecified, however are taken to be known on a given surface. The basis vectors, metric components and Jacobian  $\mathcal{J}_{\varrho\vartheta\zeta}$  of this system are detailed in appendix D.3. Note that the metric component  $g_{\varrho\vartheta}$  is in general non-zero, and so this constitutes a non-orthogonal system.

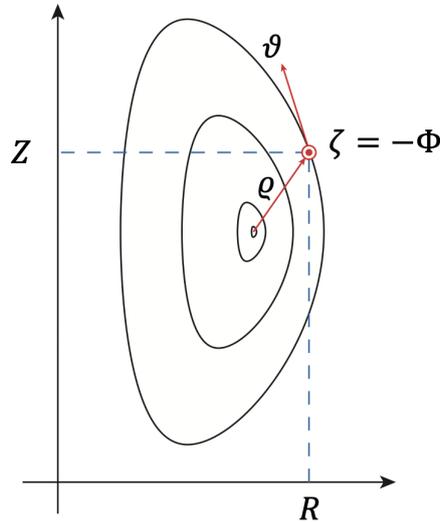


Figure 3.1: Cartoon of the cylindrical  $\{R, \Phi, Z\}$  and toroidal  $\{\varrho, \vartheta, \zeta\}$  coordinate systems. Note  $\zeta$  is directed out of the page, whereas  $\Phi$  is directed into the page.

## Magnetic field components

It is shown in appendix D.3.1 that, by requiring the magnetic field components to be periodic in  $\vartheta$ ,  $G$  in the toroidal system takes the form

$$G = \bar{c}_2(\varrho) \vartheta - \bar{c}_1(\varrho) \zeta + \tilde{G}_0(\varrho, \vartheta) \quad (3.112)$$

where  $\tilde{G}_0(\varrho, \vartheta)$  is an as-yet undetermined periodic function in  $\vartheta$ , and  $\bar{c}_2$  is a flux function. Equation 3.109 now reads

$$\begin{aligned} \mathbf{B} &= \nabla_{\varrho} \times \nabla \left[ \bar{c}_2(\varrho) \vartheta - \bar{c}_1(\varrho) \zeta + \tilde{G}_0(\varrho, \vartheta) \right] \\ &= \frac{\bar{c}_1}{\mathcal{J}_{\varrho\vartheta\zeta}} \mathbf{a}_{\vartheta} + \frac{1}{\mathcal{J}_{\varrho\vartheta\zeta}} \left( \bar{c}_2 + \frac{\partial \tilde{G}_0}{\partial \vartheta} \right) \mathbf{a}_{\zeta}. \end{aligned} \quad (3.113)$$

where we have used equation C.30. The physical interpretation of  $\bar{c}_1$  and  $\bar{c}_2$  can be found by considering the poloidal and toroidal magnetic fluxes, as shown in appendix D.3.2. One obtains

$$\bar{c}_1(\varrho) = \psi', \quad \bar{c}_2(\varrho) = \chi' \quad (3.114)$$

where  $\chi$  is the toroidal magnetic flux divided by  $2\pi$ ,  $\psi$  is the poloidal magnetic flux divided by  $2\pi$ , and here a prime denotes a radial derivative. Taking out a factor of  $\psi'$  from equation 3.113, and defining  $G_0 = \tilde{G}_0/\psi'$ , we get

$$\mathbf{B} = \psi' \nabla_{\varrho} \times \nabla [q(\varrho) \vartheta - \zeta + G_0(\varrho, \vartheta)] \quad (3.115)$$

where we have defined the safety factor

$$q(\varrho) = \frac{\chi'}{\psi'} = \frac{d\chi}{d\psi} \quad (3.116)$$

which is a flux function. Equation 3.113 hence becomes

$$\mathbf{B} = \frac{\psi'}{\mathcal{J}_{\varrho\vartheta\zeta}} \mathbf{a}_{\vartheta} + \frac{\psi'}{\mathcal{J}_{\varrho\vartheta\zeta}} \left( q(\varrho) + \frac{\partial G_0}{\partial \vartheta} \right) \mathbf{a}_{\zeta} \quad (3.117)$$

and so we see that the contravariant components of the magnetic field in the toroidal system are

$$B^{\varrho} = 0, \quad B^{\vartheta} = \frac{\psi'}{\mathcal{J}_{\varrho\vartheta\zeta}}, \quad B^{\zeta} = \frac{\psi'}{\mathcal{J}_{\varrho\vartheta\zeta}} \left( q(\varrho) + \frac{\partial G_0}{\partial \vartheta} \right). \quad (3.118)$$

The safety factor  $q$  may also be defined for tokamaks by considering the poloidal average of the ratio  $B^{\zeta}/B^{\vartheta}$ ,

$$q(\varrho) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{B^{\zeta}}{B^{\vartheta}} d\vartheta \quad (3.119)$$

where we have used the fact that  $G_0$  is periodic in  $\vartheta$ . Here we observe the relation of  $q$  to the average pitch angle of the magnetic field for a given flux surface, however we note that

these are contravariant components, not the physical magnitudes of the toroidal and poloidal fields<sup>12</sup>.

### 3.4.4 Plasma equilibrium: force balance and Ampère's law

Until this point the form of the equilibrium magnetic field has been found considering only its divergencelessness, and the definition of the magnetic fluxes. To find the form of  $G_0$ , we now include force balance, shown in appendix D.1 to be

$$\mathbf{J} \times \mathbf{B} = \nabla P \quad (3.120)$$

and the equilibrium component of Ampère's law<sup>13</sup>,

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J}. \quad (3.121)$$

#### Force balance

The covariant components of force balance are

$$\frac{\partial P}{\partial \xi^i} = \mathcal{J} \epsilon_{ijk} J^j B^k \quad (3.122)$$

using equation C.56. Writing these three equations explicitly,

$$\frac{\partial P}{\partial \varrho} = \mathcal{J}_{\varrho\vartheta\zeta} \left( J^\vartheta B^\zeta - J^\zeta B^\vartheta \right) \quad (3.123)$$

$$\frac{\partial P}{\partial \vartheta} = -\mathcal{J}_{\varrho\vartheta\zeta} J^\varrho B^\zeta \quad (3.124)$$

$$\frac{\partial P}{\partial \zeta} = \mathcal{J}_{\varrho\vartheta\zeta} J^\varrho B^\vartheta \quad (3.125)$$

where we have used  $B^\varrho = 0$ . Invoking axisymmetry to set  $\partial P / \partial \zeta = 0$  implies that  $J^\varrho = 0$ , and so akin to the magnetic field, the equilibrium current density has no component perpendicular to the flux surface. This in turn requires that the poloidal dependence of the pressure also be zero,  $\partial P / \partial \vartheta = 0$ . The equilibrium plasma pressure is therefore a flux function,  $P = P(\varrho)$ . The equation for radial force balance shall be returned to in section 3.4.5.

---

<sup>12</sup>These can be obtained from  $|\mathbf{B}_{\text{tor}}| = \sqrt{B^\zeta B_\zeta} = \sqrt{g_{\zeta\zeta} (B^\zeta)^2}$  and  $|\mathbf{B}_{\text{pol}}| = \sqrt{B^\vartheta B_\vartheta} = \sqrt{g_{\vartheta\vartheta} (B^\vartheta)^2}$ .

<sup>13</sup>Even for the case of  $\mathbf{E} \neq \mathbf{0}$  the displacement current term is negligible due to the orderings of section 3.1.2.

## Ampère's law

The contravariant components of the current density are

$$\mu_0 J^i = \frac{1}{\mathcal{J}} \epsilon^{ijk} \frac{\partial B_k}{\partial \xi^j} \quad (3.126)$$

where we have used equation C.58. Writing these explicitly,

$$\mu_0 J^\varrho = \frac{1}{\mathcal{J}_{\varrho\vartheta\zeta}} \frac{\partial B_\zeta}{\partial \vartheta} \quad (3.127)$$

$$\mu_0 J^\vartheta = -\frac{1}{\mathcal{J}_{\varrho\vartheta\zeta}} \frac{\partial B_\zeta}{\partial \varrho} \quad (3.128)$$

$$\mu_0 J^\zeta = \frac{1}{\mathcal{J}_{\varrho\vartheta\zeta}} \left( \frac{\partial B_\vartheta}{\partial \varrho} - \frac{\partial B_\varrho}{\partial \vartheta} \right) \quad (3.129)$$

where axisymmetry has been used. The condition  $J^\varrho = 0$  from force balance implies the crucial result that the covariant toroidal component of the magnetic field is a flux function,  $B_\zeta = B_\zeta(\varrho)$ , which we can use to determine  $G_0$ . Relating the contravariant and covariant components via  $B^\zeta = g^{\zeta\zeta} B_\zeta$  and using equation 3.118 one has

$$\frac{\psi'}{\mathcal{J}_{\varrho\vartheta\zeta}} \left[ q(\varrho) + \frac{\partial G_0}{\partial \vartheta} \right] = g^{\zeta\zeta} B_\zeta(\varrho) \quad (3.130)$$

which, with the requirement that  $G_0$  be periodic in  $\vartheta$  such that  $G_0(\varrho, \vartheta_0 + 2\pi) = G_0(\varrho, \vartheta_0)$ , has the solution

$$G_0(\varrho, \vartheta) = G_0(\varrho, \vartheta_0) + q(\varrho) \left[ \frac{\int_{\vartheta_0}^{\vartheta} \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} d\vartheta'}{\langle \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} \rangle_{\text{P}}} - (\vartheta - \vartheta_0) \right] \quad (3.131)$$

with

$$B_\zeta(\varrho) = \frac{q\psi'}{\langle \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} \rangle_{\text{P}}} \quad (3.132)$$

where we have introduced the notation  $\langle \dots \rangle_{\text{P}}$  to denote a poloidal average

$$\langle f \rangle_{\text{P}} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f d\vartheta. \quad (3.133)$$

The fact that  $G_0(\varrho, \vartheta_0)$  is unspecified is a manifestation of the gauge freedom in  $G$ <sup>14</sup>.

It is noted that the choices used here for the poloidal and toroidal angles  $\vartheta$  and  $\zeta$  are not unique. Different codes can make different choices regarding their definitions of poloidal and toroidal angles to alter the properties of the coordinate system, as discussed in appendix D.3.3. These include ‘straight-line’ field coordinates, also called flux coordinates [85], which make the ratio of the contravariant toroidal and poloidal components of the magnetic field a flux

<sup>14</sup>Defining an alternative potential  $G'$  via  $G = G' + \lambda(\varrho)$ , one finds  $\mathbf{B} = \nabla_\varrho \times \nabla G = \nabla_\varrho \times \nabla (G' + \lambda(\varrho)) = \nabla_\varrho \times \nabla G'$ .

function. One may also change the Jacobian to a flux function through the definition of the poloidal angle. The combination of these two is called ‘Hamada coordinates’ [87, 88]. Care must be taken therefore in comparisons between codes using different coordinate systems. For this work we shall continue using the original system  $\{\varrho, \vartheta, \zeta\}$ .

### 3.4.5 The local Grad-Shafranov solution in toroidal coordinates

Having solved for  $G_0$  we now return to radial force balance, equation 3.123. Multiplying both sides by  $\mu_0$ , and using  $P = P(\varrho)$ , it reads

$$\mu_0 \frac{dP}{d\varrho} = \mathcal{J}_{\varrho\vartheta\zeta} \left( \mu_0 J^\vartheta B^\zeta - \mu_0 J^\zeta B^\vartheta \right). \quad (3.134)$$

When combined with Ampère’s law, this can be shown to be the same as the more commonly known Grad-Shafranov equation [8], by noting that the contravariant toroidal current density in the toroidal coordinate system is closely related to that in the cylindrical system,  $J^\zeta = -J^\Phi$ , due to  $\nabla\zeta = -\nabla\Phi$ . By making this substitution one can derive the Grad-Shafranov equation, shown explicitly in appendix D.3.4:

$$\frac{\partial^2 \psi}{\partial R^2} - \frac{1}{R} \frac{\partial \psi}{\partial R} + \frac{\partial^2 \psi}{\partial Z^2} = -\mu_0 \frac{dP}{d\psi} R^2 - B_\zeta \frac{dB_\zeta}{d\psi}. \quad (3.135)$$

Note how the independent variable on the left-hand side  $\psi$  is the dependent variable on the right-hand side. This ‘mixing’ of the two coordinate systems originates from the substitution  $J^\zeta = -J^\Phi$ . To solve this for a given flux surface, an additional coordinate system is typically introduced, the Mercier-Luc system [89], and an expansion around a given surface is made, see for example [90]. In the following, the exact solution for radial force balance is presented solely in the toroidal coordinate system,  $\{\varrho, \vartheta, \zeta\}$ , bypassing the flux-surface expansion and expressing the local Grad-Shafranov condition as the solution to two simultaneous equations<sup>15</sup>.

### Force balance in toroidal coordinates

Substituting the current density components from equations 3.128 and 3.129 into 3.134 we get

$$\mu_0 \frac{dP}{d\varrho} = \left( \frac{\partial B_\varrho}{\partial \vartheta} - \frac{\partial B_\vartheta}{\partial \varrho} \right) B^\vartheta - \frac{dB_\zeta}{d\varrho} B^\zeta. \quad (3.136)$$

---

<sup>15</sup>Note that this approach is similar to that presented in appendix B of [91].

The covariant field components in our system are  $B_\vartheta = g_{\vartheta\vartheta}B^\vartheta$  and  $B_\varrho = g_{\varrho\vartheta}B^\vartheta$ , with  $B^\zeta = g^{\zeta\zeta}B_\zeta$  and  $B^\vartheta = \psi'/\mathcal{J}_{\varrho\vartheta\zeta}$ . Evaluating the partial derivatives, equation 3.136 rearranges to

$$\begin{aligned} \frac{\partial \mathcal{J}_{\varrho\vartheta\zeta}}{\partial \varrho} &= \mathcal{J}_{\varrho\vartheta\zeta} \frac{\psi''}{\psi'} - \frac{\mathcal{J}_{\varrho\vartheta\zeta}}{g_{\vartheta\vartheta}} \left( \frac{\partial g_{\varrho\vartheta}}{\partial \vartheta} - \frac{\partial g_{\vartheta\vartheta}}{\partial \varrho} - \frac{g_{\varrho\vartheta}}{\mathcal{J}_{\varrho\vartheta\zeta}} \frac{\partial \mathcal{J}_{\varrho\vartheta\zeta}}{\partial \vartheta} \right) \\ &+ \frac{(\mathcal{J}_{\varrho\vartheta\zeta})^3}{g_{\vartheta\vartheta}} \left[ \frac{\mu_0}{(\psi')^2} \frac{dP}{d\varrho} \right] + \frac{(\mathcal{J}_{\varrho\vartheta\zeta})^3 g^{\zeta\zeta}}{g_{\vartheta\vartheta}} \left[ \frac{B_\zeta}{(\psi')^2} \frac{dB_\zeta}{d\varrho} \right] \end{aligned} \quad (3.137)$$

where  $\psi'' = d^2\psi/d\varrho^2$ . The equation has been written in this form as the radial derivative of the Jacobian is the only term to contain second derivatives in  $\varrho$  of  $R$  and  $Z$ , as shown in equation D.36, and so constitutes the only  $\vartheta$ -dependent unknown provided knowledge of  $R, Z, \partial R/\partial \varrho$  and  $\partial Z/\partial \varrho$  on a given flux surface. We see therefore that radial force balance forms a constraint on the radial derivative of the Jacobian in terms of geometrical quantities, including the as-yet unknown flux function  $dB_\zeta/d\varrho$ . A second equation can be found relating these two quantities by differentiating equation 3.132 with respect to  $\varrho$ ,

$$\frac{dB_\zeta}{d\varrho} = B_\zeta \left[ \frac{q'}{q} + \frac{\psi''}{\psi'} - \frac{1}{\langle \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} \rangle_P} \left( \left\langle \frac{\partial \mathcal{J}_{\varrho\vartheta\zeta}}{\partial \varrho} g^{\zeta\zeta} \right\rangle_P + \left\langle \mathcal{J}_{\varrho\vartheta\zeta} \frac{\partial g^{\zeta\zeta}}{\partial \varrho} \right\rangle_P \right) \right]. \quad (3.138)$$

The solution to the simultaneous equations 3.137 and 3.138 can then be obtained by using equation 3.137 to calculate  $\left\langle \frac{\partial \mathcal{J}_{\varrho\vartheta\zeta}}{\partial \varrho} g^{\zeta\zeta} \right\rangle_P$  and substituting it into 3.138. This yields the solution

$$\begin{aligned} \frac{dB_\zeta}{d\varrho} &= \frac{B_\zeta}{H} \left( \frac{q'}{q} \langle \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} \rangle_P - \left\langle \mathcal{J}_{\varrho\vartheta\zeta} \frac{\partial g^{\zeta\zeta}}{\partial \varrho} \right\rangle_P - \left[ \frac{\mu_0}{(\psi')^2} \frac{dP}{d\varrho} \right] \left\langle \frac{(\mathcal{J}_{\varrho\vartheta\zeta})^3 g^{\zeta\zeta}}{g_{\vartheta\vartheta}} \right\rangle_P \right. \\ &\quad \left. + \left\langle \frac{\mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta}}{g_{\vartheta\vartheta}} \left( \frac{\partial g_{\varrho\vartheta}}{\partial \vartheta} - \frac{\partial g_{\vartheta\vartheta}}{\partial \varrho} - \frac{g_{\varrho\vartheta}}{\mathcal{J}_{\varrho\vartheta\zeta}} \frac{\partial \mathcal{J}_{\varrho\vartheta\zeta}}{\partial \vartheta} \right) \right\rangle_P \right) \end{aligned} \quad (3.139)$$

where

$$H = \langle \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} \rangle_P + \left( \frac{B_\zeta}{\psi'} \right)^2 \left\langle \frac{(\mathcal{J}_{\varrho\vartheta\zeta})^3 (g^{\zeta\zeta})^2}{g_{\vartheta\vartheta}} \right\rangle_P \quad (3.140)$$

which is the condition to satisfy force balance on a given surface. Note the cancellation of  $\psi''$ . The radial derivative of the Jacobian can then be obtained by substituting 3.139 back into equation 3.137.

### Formalism comparison

By comparing the calculation of the geometric quantity  $g^{\alpha\alpha}$  (equation D.97)<sup>16</sup> using the results of this section with those obtained from the CGYRO formalism [90] for a selection of tokamak equilibria<sup>17</sup>, the equivalence of the two approaches to the local Grad Shafranov solution can be demonstrated. This is shown in figure 3.2.

<sup>16</sup>This is a metric term in the upcoming field-aligned system, which depends on the radial derivatives given by equations 3.139 and 3.137.

<sup>17</sup>These equilibria are detailed in section 4.6

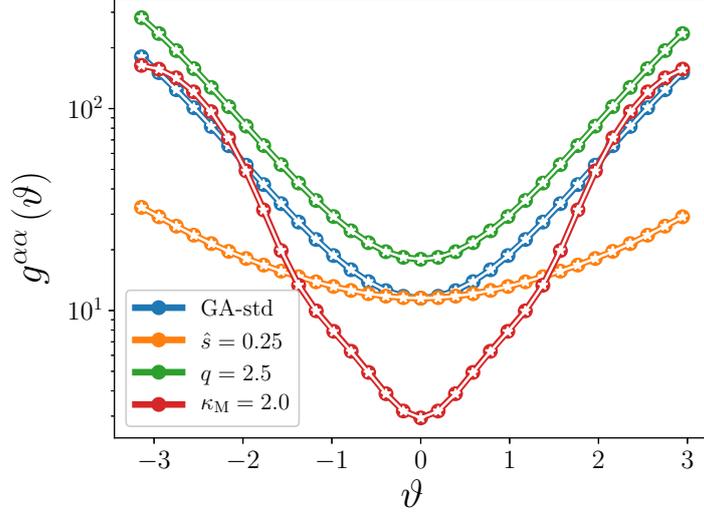


Figure 3.2: Comparison of the quantity  $g^{\alpha\alpha}$  obtained from the Mercier-Luc formalism using CGYRO (coloured circles), and that obtained in this section (white stars) for a selection of tokamak equilibria. Provided the other terms in the equation are known,  $g^{\alpha\alpha}$  becomes a proxy for agreement between the radial derivatives in the system. For all equilibria the curves overlap everywhere, indicating the equivalence of the two approaches.

### 3.4.6 Plasma shape

The analysis performed in this section so far has solved the divergenceless condition of the magnetic field, force balance and Ampère’s law for a given flux surface  $\varrho$ , solving for the  $\vartheta$  dependence of the equilibrium. This section has said nothing of integrating the equilibrium radially, which in general must be done numerically and requires knowledge of radial profiles, such as the pressure  $P(\varrho)$  and safety factor  $q(\varrho)$ . What this analysis therefore constitutes is a boundary condition for such an integration<sup>18</sup>.

The local geometry solution has been calculated assuming that the shape functions,  $R$ ,  $Z$ ,  $\partial R/\partial\varrho$  and  $\partial Z/\partial\varrho$  have been known on a given flux surface. They are subject to constraints: they must all be real, periodic functions in  $\vartheta$ , and  $R$  and  $\mathcal{J}_{\varrho\vartheta\zeta}$  (equation D.35) must be positive everywhere. Provided that these conditions are satisfied, then we are free to define them, with the most general specification of the shape functions being through the use of Fourier series. More frequently a surface shape parameterisation will be used, reducing the freedom in the shape function specification to a few more interpretable parameters.

<sup>18</sup>As this work focuses on local equilibria, those defined on a single surface, radial integration is not considered in this work.

### Miller parameterisation

A common flux-surface parameterisation choice, and one used in this work, is the Miller parameterisation. First used in [92], additional shaping parameters were later introduced in [93], and is defined on a surface  $\varrho = \varrho_0$  to be<sup>19</sup>

$$R(\varrho_0, \vartheta) = R_{0M} + r_M \cos(\vartheta + \arcsin[\delta_M] \sin(\vartheta)) \quad (3.141)$$

$$Z(\varrho_0, \vartheta) = Z_{0M} + \kappa_M r_M \sin(\vartheta + \zeta_M \sin(2\vartheta)). \quad (3.142)$$

One observes that this parameterisation makes  $R(\varrho_0, \vartheta)$  an even function of  $\vartheta$ , and  $Z(\varrho_0, \vartheta)$  an odd function around the elevation  $Z_{0M}$ , known as up-down symmetric,  $Z(\varrho_0, \vartheta) - Z_{0M} = -[Z(\varrho_0, -\vartheta) - Z_{0M}]$ .

Here  $R_{0M} = \frac{1}{2}(R(\varrho_0, 0) + R(\varrho_0, -\pi))$  marks the centre of the surface at the midplane, and analogously for the elevation, we have  $Z_{0M} = \frac{1}{2}(Z(\varrho_0, \frac{\pi}{2}) + Z(\varrho_0, -\frac{\pi}{2}))$ . The quantity  $r_M = \frac{1}{2}(R(\varrho_0, 0) - R(\varrho_0, -\pi))$  characterises the span of the surface at the midplane, and the elongation  $\kappa_M$  describes the maximum of the surface in  $Z$  relative to its maximum in  $R$ ,  $\kappa_M = (Z(\varrho_0, \frac{\pi}{2}) - Z_{0M}) / (R(\varrho_0, 0) - R_{0M})$ . The ‘triangularity’ of the surface,  $\delta_M$ , can be interpreted via  $R(\varrho_0, \frac{\pi}{2}) = R_{0M} - \delta_M r_M$ , with the ‘squareness’  $\zeta_M$  less readily defined, however its effect on the flux-surface shape, along with those of the other parameters, is demonstrated in figure 3.3.

We now consider the parameterisation of the radial derivatives,  $\partial R / \partial \varrho|_{\varrho_0, \vartheta}$  and  $\partial Z / \partial \varrho|_{\varrho_0, \vartheta}$ . To define these parameterisations, one assumes that a closely adjacent surface  $\varrho = \varrho_0 + \Delta\varrho$  can also be parameterised using equations 3.141 and 3.142. Taking the limit  $\Delta\varrho \rightarrow 0$ , this amounts to an assumption of radial dependence on the shape parameters and differentiating, which gives

$$\left. \frac{\partial R}{\partial \varrho} \right|_{\varrho_0, \vartheta} = \left. \frac{dR_0}{d\varrho} \right|_{\varrho_0} + \cos(\vartheta + \arcsin[\delta_M] \sin(\vartheta)) - \frac{r_M \frac{d\delta_M}{d\varrho}}{\sqrt{1 - \delta_M^2}} \sin(\vartheta) \sin(\vartheta + \arcsin[\delta_M] \sin(\vartheta)) \quad (3.143)$$

$$\left. \frac{\partial Z}{\partial \varrho} \right|_{\varrho_0, \vartheta} = \left. \frac{dZ_0}{d\varrho} \right|_{\varrho_0} + \left[ \frac{d\kappa_M}{d\varrho} r_M + \kappa_M \right] \sin(\vartheta + \zeta_M \sin(2\vartheta)) + \kappa_M r_M \frac{d\zeta_M}{d\varrho} \sin(2\vartheta) \cos(\vartheta + \zeta_M \sin(2\vartheta)). \quad (3.144)$$

A description of the four shape functions with the Miller geometry therefore requires values for 11 parameters,  $\{r_M, \kappa_M, d\kappa_M/d\varrho, \delta_M, d\delta_M/d\varrho, R_0, dR_0/d\varrho, Z_0, dZ_0/d\varrho, \zeta_M, d\zeta_M/d\varrho\}$ . These derivative parameters do not affect the shape of the given flux surface, however do influence the geometry through the metric terms.

---

<sup>19</sup>Note, the  $\arcsin[\delta_M]$  term is not used consistently by all authors.

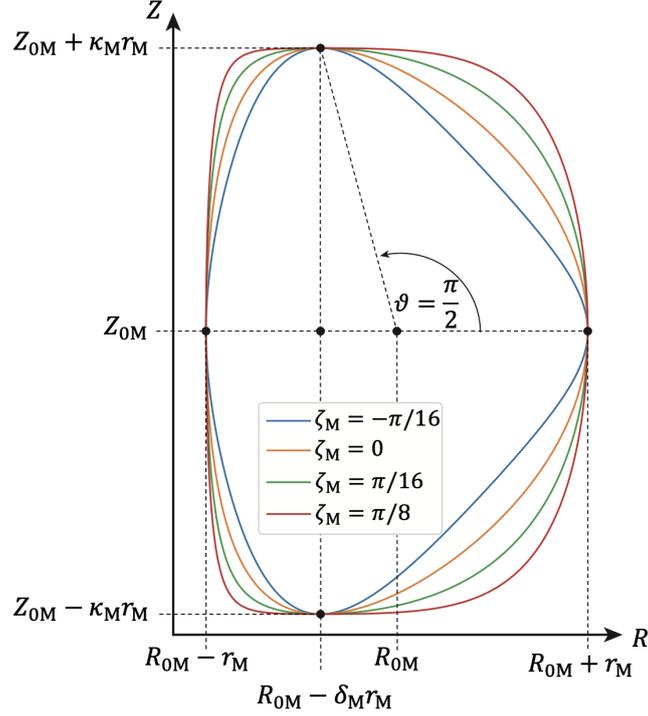


Figure 3.3: Examples of flux surfaces constructed using the Miller parameterisation with various anchor points marked. Note that the poloidal angle does not necessarily correspond to the parameterisation angle.

### 3.4.7 Toroidal coordinate summary

The equilibrium magnetic field may be written

$$\begin{aligned} \mathbf{B} &= \psi' \nabla \varrho \times \nabla (q(\varrho) \vartheta - \zeta + G_0(\varrho, \vartheta)) \\ &= \frac{\psi'}{\mathcal{J}_{\varrho\vartheta\zeta}} \mathbf{a}_\vartheta + g^{\zeta\zeta} B_\zeta(\varrho) \mathbf{a}_\zeta \end{aligned} \quad (3.145)$$

where  $G_0$  is a periodic function given by 3.131 and  $B_\zeta$  is a flux function given by 3.132. The equilibrium plasma pressure is a flux function,  $P = P(\varrho)$ , such that

$$\nabla P = \frac{dP}{d\varrho} \nabla \varrho. \quad (3.146)$$

Using equations 3.128 and 3.134, the current density is

$$\mu_0 \mathbf{J} = -\frac{1}{\mathcal{J}_{\varrho\vartheta\zeta}} \frac{dB_\zeta}{d\varrho} \mathbf{a}_\vartheta - \frac{1}{\psi'} \left( B_\zeta \frac{dB_\zeta}{d\varrho} + \mu_0 \frac{dP}{d\varrho} \right) \mathbf{a}_\zeta \quad (3.147)$$

where  $dB_\zeta/d\varrho$  is given by equation 3.139. These expressions solve the equations 3.107, 3.120 and 3.121 on a given flux surface under the assumption of axisymmetry, and are defined in terms of:

- The safety factor  $q$  and its radial derivative,  $q'$ , often parameterised using the dimensionless quantity *magnetic shear*  $\hat{s} = (\varrho_0/q) q'$ .
- The equilibrium plasma pressure,  $dP/d\varrho$ .
- The four shape functions,  $R(\varrho_0, \vartheta)$ ,  $Z(\varrho_0, \vartheta)$ ,  $\partial R/\partial \varrho(\varrho_0, \vartheta)$  and  $\partial Z/\partial \varrho(\varrho_0, \vartheta)$ , typically defined through a parameterisation like Miller.

### 3.4.8 Global field-aligned coordinates

Having completely described the tokamak geometry of a given surface, we have now achieved our goal of finding the surface potentials that describe the equilibrium magnetic field, namely

$$\mathbf{B} = \psi' \nabla \varrho \times \nabla \alpha. \quad (3.148)$$

where  $\alpha = q(\varrho) \vartheta - \zeta + G_0(\varrho, \vartheta)$ . By transforming to coordinates  $r = \varrho$ ,  $\alpha = q(\varrho) \vartheta - \zeta + G_0(\varrho, \vartheta)$ , and an arbitrary third coordinate, we obtain our field-aligned system. In appendix D.4.1 it is shown that making the choice for the third coordinate, which we label  $\theta$ , to be  $\theta = \vartheta$  has several useful properties, namely that axisymmetry is expressed solely in the  $\alpha$  coordinate  $\partial/\partial \zeta = -\partial/\partial \alpha$ , the Jacobian remains unchanged over the transformation  $\mathcal{J}_{\varrho\vartheta\zeta} = \mathcal{J}_{r\alpha\theta}$  and our flux-surface shape functions remain unchanged, e.g.  $R(\varrho_0, \vartheta) \rightarrow R(r_0, \theta)$ .

We transform from the toroidal system to the field-aligned system,  $\{\varrho, \vartheta, \zeta\} \rightarrow \{r, \alpha, \theta\}$ , defined by

$$r = \varrho \qquad \qquad \varrho = r \quad (3.149)$$

$$\alpha = q(\varrho) \vartheta - \zeta + G_0(\varrho, \vartheta) \qquad \vartheta = \theta \quad (3.150)$$

$$\theta = \vartheta \qquad \qquad \zeta = q(r) \theta - \alpha + G_0(r, \theta) \quad (3.151)$$

where geometric quantities of interest such as basis vectors<sup>20</sup> and metric components are given in appendix D.4.2. A graphical representation of this coordinate system is shown in figure 3.4. The equilibrium magnetic field becomes

$$\mathbf{B} = \frac{\psi'}{\mathcal{J}_{r\alpha\theta}} \mathbf{a}_\theta \quad (3.152)$$

with the current density

$$\mu_0 \mathbf{J} = \frac{\mu_0}{\psi'} \frac{dP}{dr} \mathbf{a}_\alpha - \frac{1}{\mathcal{J}_{r\alpha\theta}} \frac{dB_\zeta}{dr} \mathbf{a}_\theta \quad (3.153)$$

and pressure gradient

$$\nabla P = \frac{dP}{dr} \nabla r. \quad (3.154)$$

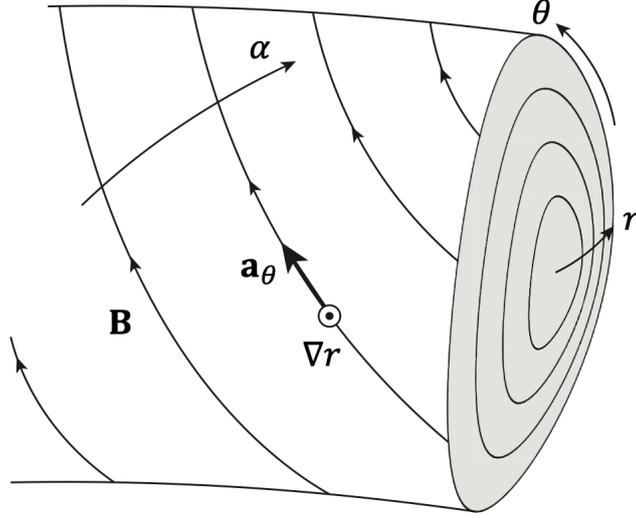


Figure 3.4: Schematic of the field-aligned system  $\{r, \alpha, \theta\}$ , including the alignment of the equilibrium magnetic field  $\mathbf{B}$  with the covariant basis vector of the  $\theta$  coordinate,  $\mathbf{a}_\theta$ .

At constant  $\{r, \alpha\}$ ,  $\theta$  measures the distance along the magnetic field, whereas at constant  $\{r, \zeta\}$ ,  $\theta$  measures the distance along a poloidal turn. We now consider properties of the new coordinate system.

### Global periodicity

Physical periodicity corresponds to a function  $f$  being unchanged for  $M$  poloidal turns at constant toroidal angle, and  $N$  toroidal turns at constant poloidal angle, for integers  $M$  and  $N$ . In the toroidal coordinate system, this corresponds to

$$f(\varrho, \vartheta + 2M\pi, \zeta) = f(\varrho, \vartheta, \zeta) \quad (3.155)$$

$$f(\varrho, \vartheta, \zeta + 2N\pi) = f(\varrho, \vartheta, \zeta). \quad (3.156)$$

Now expressed in the field-aligned system, we have for poloidal periodicity

$$f(r(\varrho), \alpha(\varrho, \vartheta + 2M\pi, \zeta), \theta(\vartheta + 2M\pi)) = f(r, \alpha + 2M\pi q(r), \theta + 2M\pi) \quad (3.157)$$

and thus

$$f(r, \alpha, \theta) = f(r, \alpha + 2M\pi q(r), \theta + 2M\pi). \quad (3.158)$$

Correspondingly for toroidal periodicity,

$$f(r(\varrho), \alpha(\varrho, \vartheta, \zeta + 2N\pi), \theta(\vartheta)) = f(r, \alpha - 2N\pi, \theta) \quad (3.159)$$

<sup>20</sup>Note that although for example  $\theta = \vartheta$ , one finds that  $\mathbf{a}_\theta \neq \mathbf{a}_\vartheta$ , motivating the use of distinct symbols.

$$f(r, \alpha, \theta) = f(r, \alpha - 2N\pi, \theta). \quad (3.160)$$

The toroidal periodicity condition allows an arbitrary function to be expressed as a Fourier series in  $\alpha$ ,

$$f(r, \alpha, \theta) = \sum_{k_\alpha} \hat{f}_{k_\alpha}(r, \theta) e^{ik_\alpha \alpha} \quad (3.161)$$

where  $k_\alpha = N_\alpha n_\alpha$  for all integers  $n_\alpha$ , and  $N_\alpha$  is a single integer that determines the size of the  $\alpha$  domain considered [94]. Applying the poloidal periodicity condition 3.158 to equation 3.161, one finds

$$\sum_{k_\alpha} \hat{f}_{k_\alpha}(r, \theta) e^{ik_\alpha \alpha} = \sum_{k_\alpha} \hat{f}_{k_\alpha}(r, \theta + 2M\pi) e^{ik_\alpha \alpha} e^{ik_\alpha 2M\pi q(r)} \quad (3.162)$$

which implies the  $k_\alpha$  Fourier amplitudes must satisfy

$$\hat{f}_{k_\alpha}(r, \theta) = \hat{f}_{k_\alpha}(r, \theta + 2M\pi) e^{ik_\alpha 2M\pi q(r)}. \quad (3.163)$$

Due to the presence of  $q(r)$  in the exponent of the right-hand side, this condition generally prevents either the  $r$  coordinate or the  $\theta$  coordinate from being expressed as a Fourier series when considering the coordinate system globally. This condition shall be revisited when examining the local limit.

### Geometric operators in the field-aligned system

One of the reasons for deriving the field-aligned system is the relegation of the parallel dynamics to a single coordinate, separating the disparate turbulent scales parallel and perpendicular to the magnetic field. For the gradient of a function  $f$ , we have

$$\nabla f = \frac{\partial f}{\partial r} \nabla r + \frac{\partial f}{\partial \alpha} \nabla \alpha + \frac{\partial f}{\partial \theta} \nabla \theta. \quad (3.164)$$

Using equation 3.152, the magnitude of the equilibrium magnetic field is  $|\mathbf{B}| = \psi' \sqrt{g_{\theta\theta}} / \mathcal{J}_{r\alpha\theta}$ , which makes the unit vector of the magnetic field in the system

$$\hat{\mathbf{b}} = \frac{1}{\sqrt{g_{\theta\theta}}} \mathbf{a}_\theta. \quad (3.165)$$

The parallel derivative in the field-aligned system is therefore

$$\hat{\mathbf{b}} \cdot \nabla f = \nabla_{\parallel} f = \frac{1}{\sqrt{g_{\theta\theta}}} \frac{\partial f}{\partial \theta} \quad (3.166)$$

confirming its isolation to the  $\theta$  coordinate. The perpendicular derivative, using  $\mathbf{a}_\theta = g_{r\theta} \nabla r + g_{\alpha\theta} \nabla \alpha + g_{\theta\theta} \nabla \theta$ , is

$$\begin{aligned} \nabla_{\perp} f &= \nabla f - \nabla_{\parallel} f \hat{\mathbf{b}} = \left( \frac{\partial f}{\partial r} - \frac{g_{r\theta}}{g_{\theta\theta}} \frac{\partial f}{\partial \theta} \right) \nabla r + \left( \frac{\partial f}{\partial \alpha} - \frac{g_{\alpha\theta}}{g_{\theta\theta}} \frac{\partial f}{\partial \theta} \right) \nabla \alpha \\ &= \frac{\partial f}{\partial r} \nabla r + \frac{\partial f}{\partial \alpha} \nabla \alpha + \mathcal{O}(\rho_* k_{\perp} f) \end{aligned} \quad (3.167)$$

and so for fluctuating quantities, the two perpendicular-to-field coordinates  $r$  and  $\alpha$  capture the perpendicular derivatives to lowest order.

### Orthonormal basis vectors

In the derivation of gyrokinetics we used an orthonormal basis,  $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{b}}\}$ . We can now consider such a basis explicitly. The unit magnetic field  $\hat{\mathbf{b}}$  is given by equation 3.165, which we know to be orthogonal to  $\nabla r$  and  $\nabla\alpha$  everywhere. Due to the importance of the radial direction in our analysis, we define  $\hat{\mathbf{e}}_1$  to be a unit vector in the radial direction

$$\hat{\mathbf{e}}_1 = \frac{1}{\sqrt{g^{rr}}} \nabla r \quad (3.168)$$

which leaves the final vector perpendicular to both, defining the binormal direction, to be obtained from  $\hat{\mathbf{e}}_2 = \hat{\mathbf{b}} \times \hat{\mathbf{e}}_1$ . This is

$$\begin{aligned} \hat{\mathbf{e}}_2 &= \frac{1}{\sqrt{g^{rr}g_{\theta\theta}}} \mathbf{a}_\theta \times \nabla r \\ &= \frac{1}{\sqrt{g^{rr}g_{\theta\theta}}} (g_{\alpha\theta} \nabla\alpha \times \nabla r + g_{\theta\theta} \nabla\theta \times \nabla r) \\ &= \frac{1}{\mathcal{J}_{r\alpha\theta} \sqrt{g^{rr}g_{\theta\theta}}} (g_{\theta\theta} \mathbf{a}_\alpha - g_{\alpha\theta} \mathbf{a}_\theta) \\ &= \frac{g_{\theta\theta} \mathbf{a}_\alpha - g_{\alpha\theta} \mathbf{a}_\theta}{\sqrt{g_{\theta\theta}} \sqrt{g_{\alpha\alpha}g_{\theta\theta} - (g_{\alpha\theta})^2}} \end{aligned} \quad (3.169)$$

where we have used equation C.47. Taking components of a vector in this basis makes the components physical, in the sense that they have the same dimensions as the original vector, which is not generally true for contravariant and covariant components. For a wavevector  $\mathbf{k} = k_r \nabla r + k_\alpha \nabla\alpha + k_\theta \nabla\theta = k_1 \hat{\mathbf{e}}_1 + k_2 \hat{\mathbf{e}}_2 + k_\parallel \hat{\mathbf{b}}$  for example, we have

$$k_1 = \frac{1}{\sqrt{g^{rr}}} k^r, \quad k_2 = \frac{g_{\theta\theta} k_\alpha - g_{\alpha\theta} k_\theta}{\sqrt{g_{\theta\theta}} \sqrt{g_{\alpha\alpha}g_{\theta\theta} - (g_{\alpha\theta})^2}}, \quad k_\parallel = \frac{k_\theta}{\sqrt{g_{\theta\theta}}} \quad (3.170)$$

which are distinct from the covariant components, with the dependence on metric terms causing these components to vary with  $\theta$ .

### 3.4.9 The local limit

We shall now consider focusing our domain around a single equilibrium magnetic field line, on a perpendicular scale  $L$  much larger than the size of the turbulent fluctuations  $\mathcal{O}(\rho_i)$  but small compared with the macroscopic length scale  $a$ , such that  $\rho_i \ll L \ll a$ , while maintaining the extent of the parallel direction. This is the local limit, and a graphical representation of this domain is shown in figure 3.5.

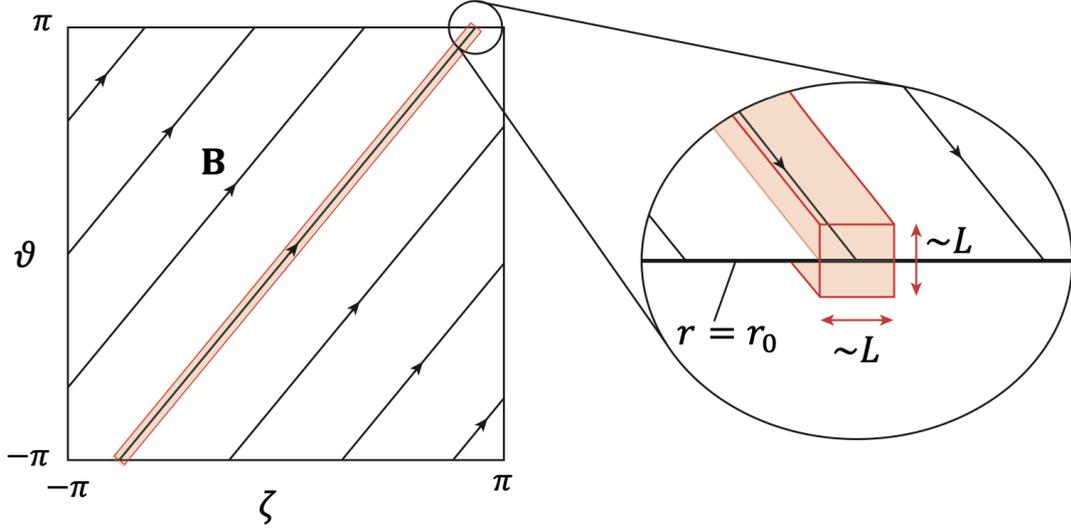


Figure 3.5: Schematic of the local limit domain of the ‘flux tube’, centered around a single equilibrium field line. Note the flattening of the tokamak flux surface is for illustrative purposes only, and the lines have been displayed as straight for simplicity, not as a consequence of flux coordinates.

The local limit causes equilibrium quantities to be approximately constant perpendicular across the domain, and allows for fluctuating quantities to be able to expressed as a Fourier series in the radial direction. Because of the independence of equilibrium quantities from  $\alpha$  due to axisymmetry, average turbulence quantities calculated in the local limit apply across the entire flux surface,  $r_0$ . This therefore represents the most efficient volume possible to obtain the turbulence characteristics for a given flux surface.

### Equilibrium radial expansion

To evaluate the local limit we first define a subsidiary perturbation parameter  $\epsilon_* = L/a$ . We then expand equilibrium quantities in  $\epsilon_*$  and take the limit  $\epsilon_* \rightarrow 0$  while retaining  $\rho_* \rightarrow 0$  for finite  $\rho_*/\epsilon_* = \rho_i/L$ . For the equilibrium temperature of a species  $s$  for example, expanding around a surface  $r = r_0$  we have<sup>21</sup>

$$T_{s,0}(r) = T_{s,0}(r_0) + (r - r_0) \left. \frac{dT_{s,0}}{dr} \right|_{r=r_0} + \dots = T_{s,0}(r_0) + \mathcal{O}(\epsilon_* T_{s,0}). \quad (3.171)$$

and so in the limit  $\epsilon_* \rightarrow 0$  equilibrium quantities are constant across the box.

<sup>21</sup>Equilibrium quantities relating to the Maxwellian background distribution are independent of  $\theta$  due to equation 3.93.

## Binormal domain

A consideration of the value of the integer  $N_\alpha$  can now be made in the local limit. Requiring that our binormal domain size is physically of order  $L$ , the smallest non-zero binormal wavevector should have a wavelength of this order. To obtain the physical length we use  $\hat{\mathbf{e}}_2$  (equation 3.169). Taking the dot product of  $\hat{\mathbf{e}}_2$  with the wavevector  $\mathbf{k}$ , then to lowest order  $k_2$  evaluated for  $n_\alpha = 1$  is

$$k_2 = \frac{g_{\theta\theta}N_\alpha}{\sqrt{g_{\theta\theta}}\sqrt{g_{\alpha\alpha}g_{\theta\theta} - (g_{\alpha\theta})^2}} \quad (3.172)$$

which we require to be  $\mathcal{O}(1/L)$ . The metric components  $g_{\alpha\alpha}$ ,  $g_{\theta\theta}$  and  $g_{\alpha\theta}$  are all of order  $\mathcal{O}(a^2)$ , and thus we find

$$N_\alpha = \mathcal{O}\left(\frac{a}{L}\right) = \mathcal{O}\left(\frac{1}{\epsilon_*}\right). \quad (3.173)$$

In the local limit therefore  $N_\alpha$  becomes an arbitrarily large integer, which physically corresponds to focusing the binormal domain around a single equilibrium field line.

## Radial Fourier representation and local periodicity

The local limit allows the radial dependence of fluctuating quantities to be expressed as a Fourier series, whereas in the global case (equation 3.163) it cannot. For convenience, we introduce the radial coordinate  $x$ , such that

$$x = r - r_0. \quad (3.174)$$

Analysing the argument of the exponent in equation 3.163, we radially expand the safety factor. The argument of the exponent is, specifying the orders,

$$\begin{aligned} iN_\alpha n_\alpha 2\pi M q(r) &= \underbrace{iN_\alpha n_\alpha 2\pi M q(r_0)}_{\mathcal{O}(1/\epsilon_*)} + \underbrace{xiN_\alpha n_\alpha 2\pi M \frac{dq}{dr}\Big|_{r_0}}_{\mathcal{O}(1)} + \\ &\quad \underbrace{\frac{1}{2}x^2 iN_\alpha n_\alpha 2\pi M \frac{d^2q}{dr^2}\Big|_{r_0}}_{\mathcal{O}(\epsilon_*)} + \mathcal{O}(\epsilon_*^2) \end{aligned} \quad (3.175)$$

The first term is of order  $1/\epsilon_*$ , due to  $N_\alpha$ . The second term becomes of order 1, as the  $1/\epsilon_*$  ordering of  $N_\alpha$  is balanced by the  $\epsilon_*$  ordering of  $x dq/dr|_{r_0}$ , and the third term is of order  $\epsilon_*$ . In the limit  $\epsilon_* \rightarrow 0$ , the first term when exponentiated becomes

$$\lim_{\epsilon_* \rightarrow 0} e^{iN_\alpha n_\alpha 2\pi M q(r_0)} \rightarrow 1 \quad (3.176)$$

as the exponent can be assumed to be an arbitrarily large integer multiple of  $2\pi^{22}$ . For the third term and beyond the exponentiated terms just becomes 1 as the exponent tends to zero,  $e^0 = 1$ . Only the second term is retained in the local limit, such that equation 3.163 becomes

$$\hat{f}_{k_\alpha}(x, \theta) = \hat{f}_{k_\alpha}(x, \theta + 2M\pi) e^{xi k_\alpha 2M\pi \frac{dq}{dr} \Big|_{r_0}}. \quad (3.177)$$

This now allows the radial direction to be expressed as a Fourier series, with a non-trivial periodicity condition. Expanding equation 3.177 in  $x$ ,

$$\sum_{k_x} \hat{f}_{k_x, k_\alpha}(\theta) e^{ik_x x} = \sum_{k_x} \hat{f}_{k_x, k_\alpha}(\theta + 2M\pi) e^{ik_x x} e^{xi k_\alpha 2M\pi \frac{dq}{dr} \Big|_{r_0}} \quad (3.178)$$

where  $k_x = 2\pi n_x / L_x$ , for integer  $n_x$  and radial domain length  $L_x$ . Multiplying both sides by  $e^{-ik'_x x}$  and averaging over the  $k_x$  domain, we have

$$\hat{f}_{k'_x, k_\alpha}(\theta) = \sum_{k_x} \hat{f}_{k_x, k_\alpha}(\theta + 2M\pi) \frac{1}{L_x} \int_{x_0}^{x_0 + L_x} e^{i(k_x + k_\alpha 2M\pi \frac{dq}{dr} \Big|_{r_0} - k'_x)x} dx \quad (3.179)$$

where  $x_0$  is the minimum value of the  $x$  domain. The average on the right-hand side isolates a mode under the condition that  $k_\alpha 2M\pi \frac{dq}{dr} \Big|_{r_0} = 2\pi p / L_x$ , for integer  $p$ . This puts a quantisation condition on the radial domain length  $L_x$ , which can be shown to be

$$L_x = \frac{w}{N_\alpha \frac{dq}{dr} \Big|_{r_0}} \quad (3.180)$$

for order 1 integer  $w$ , dubbed the ‘box size’ integer. The wavenumber picked out by the averaging is therefore  $k_x = k'_x - 2\pi(Mn_\alpha w) / L_x$ . This completes the derivation of the poloidal periodicity condition for the 2D Fourier amplitudes in the local limit, given by

$$\hat{f}_{k_x, k_\alpha}(\theta) = \hat{f}_{k_x - \frac{2\pi}{L_x}(Mn_\alpha w), k_\alpha}(\theta + 2M\pi). \quad (3.181)$$

### Binormal coordinate redefinition

Fundamentally,  $\alpha$  is an angular coordinate. However, it is desirable to recast it as a coordinate with dimensions of length, so as to make comparisons to the other length scales in the system. We do this by introducing the coordinate  $y$ , such that

$$y = C_y \alpha \quad (3.182)$$

where  $C_y$  is an equilibrium quantity with dimensions of length. This transformation redefines the binormal wavenumber  $k_y = k_\alpha / C_y$ , via

$$k_y = \frac{N_\alpha n_\alpha}{C_y} = \frac{2\pi n_y}{L_y} \quad (3.183)$$

<sup>22</sup>Here we have assumed  $\lim_{N_\alpha \rightarrow \infty} N_\alpha q(r_0)$  is an integer.

where the set of integers  $n_\alpha$  has been relabelled to  $n_y$ , and  $L_y$  is the ‘binormal domain length’,

$$L_y = \frac{2\pi C_y}{N_\alpha}. \quad (3.184)$$

It is here that different definitions of the binormal wavenumber can arise between codes, as one can make different choices for  $C_y$ . CGYRO [38] for example chooses  $C_{y,\text{CGYRO}} = r_0/q(r_0)$ , and GKW [35], in the notation of this work, makes the choice  $C_{y,\text{GKW}} = 1/\sqrt{g^{\alpha\alpha}(\theta=0)}$ .

Note that because  $L_y$  is proportional to  $C_y$  which, other than being an equilibrium quantity with dimensions of length, is arbitrary,  $L_y$  is also in a sense arbitrary, and it is advised to not use this quantity as a measure of comparing binormal domain resolutions. One could for example define  $C'_y = 2C_y$ , and quoting  $L_y$  it would appear that the binormal domain is twice as large, whereas this simply arises from definitions in  $C_y$ . By comparing the minimum  $k_y$  ( $n_y = 1$ ) to a reference gyroradius  $\rho_{\text{ref}}$  and using a reference macroscopic length scale  $a_{\text{ref}}$ , we find that fundamentally the quantity that can be compared between codes is the product  $N_\alpha \rho_{*,\text{ref}}$ , where  $\rho_{*,\text{ref}} = \rho_{\text{ref}}/a_{\text{ref}}$ ,

$$(k_y \rho_{\text{ref}})_{\text{min}} = \frac{N_\alpha \rho_{*,\text{ref}}}{(C_y/a_{\text{ref}})}. \quad (3.185)$$

Often in nonlinear gyrokinetic simulations one sets the minimum value of  $k_y \rho_{\text{ref}}$ , which we see is a proxy for  $N_\alpha \rho_{*,\text{ref}}$ . Given the orderings of  $N_\alpha \rightarrow \infty$ ,  $\rho_* \rightarrow 0$ , this is a quantity that characterises the size of the intermediate length scale with respect to the gyroradius,  $\sim \rho_i/L$ .

### Local wavenumbers

From the quantisation condition of  $L_x$ , equation 3.180, one may show for the minimum radial wavenumber in the domain

$$\begin{aligned} (k_x \rho_{\text{ref}})_{\text{min}} &= \frac{2\pi}{L_x} \rho_{\text{ref}} \\ &= (N_\alpha \rho_{*,\text{ref}}) \frac{2\pi}{w} \left[ a_{\text{ref}} \frac{dq}{dr} \Big|_{r_0} \right] \\ &= \frac{2\pi C_y}{w} \frac{dq}{dr} \Big|_{r_0} (k_y \rho_{\text{ref}})_{\text{min}}. \end{aligned} \quad (3.186)$$

This explicitly demonstrates the effect of the box size integer  $w$ , which allows one to alter the radial domain for a constant binormal domain. In this work, we use the  $C_y$  chosen by CGYRO,  $C_y = r_0/q(r_0)$ , such that

$$(k_y \rho_{\text{ref}})_{\text{min}} = q(r_0) \frac{N_\alpha \rho_{*,\text{ref}}}{(r_0/a_{\text{ref}})} \quad (3.187)$$

$$(k_x \rho_{\text{ref}})_{\text{min}} = \left[ \frac{2\pi \hat{s}}{w} \right] (k_y \rho_{\text{ref}})_{\text{min}}. \quad (3.188)$$

Our final transformation to the ‘local field-aligned system’ is therefore  $\{r, \alpha, \theta\} \rightarrow \{x, y, \theta\}$ <sup>23</sup>, where the geometric quantities of interest are shown in appendix D.5.1. Neither the covariant nor the contravariant basis vector for the third coordinate is changed for this transformation (and hence metric terms), and so one can unambiguously retain the label  $\theta$ . In this system, the magnetic field is expressed

$$\begin{aligned} \mathbf{B} &= B_{\text{unit}} \nabla x \times \nabla y \\ &= \frac{B_{\text{unit}}}{\mathcal{J}_{xy\theta}} \mathbf{a}_\theta \end{aligned} \quad (3.189)$$

where we have defined [95]

$$B_{\text{unit}} = \frac{q(r_0)}{r_0} \psi'(r_0). \quad (3.190)$$

Bringing the foregoing together, in the local limit a fluctuating quantity can be written as a Fourier series in radial and binormal coordinates  $x$  and  $y$ ,

$$f(x, y, \theta, t) = \sum_{k_x} \sum_{k_y} \hat{f}_{k_x, k_y}(\theta, t) e^{ik_x x} e^{ik_y y} \quad (3.191)$$

where the Fourier amplitudes are subject to the periodicity condition given by equation 3.181, now expressed using the  $y$  coordinate as

$$\hat{f}_{k_x + \frac{2\pi}{L_x}(Mn_y w), k_y}(\theta) = \hat{f}_{k_x, k_y}(\theta + 2M\pi). \quad (3.192)$$

### 3.5 Local gyrokinetics

The local-field aligned geometry is now applied to the gyrokinetic equation. The terms are derived explicitly in appendix D.5.2, for which it is shown equation 3.97 becomes

$$\begin{aligned} \frac{\partial h_s}{\partial t} + \frac{U_{0,\parallel}}{\sqrt{g_{\theta\theta}}} \frac{\partial h_s}{\partial \theta} + \mathbf{v}_D \cdot \frac{\partial h_s}{\partial \mathbf{R}} + \frac{1}{B_{\text{unit}}} \delta \left[ \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial x} \frac{\partial h_s}{\partial y} - \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial y} \frac{\partial h_s}{\partial x} \right] = \\ \frac{q_s F_{s,0}}{T_{s,0}} \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial t} - \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial y} \frac{F_{s,0}}{B_{\text{unit}}} \left[ \frac{1}{L_{n_s}} + \frac{1}{L_{T_s}} \left( \frac{\varepsilon}{T_{s,0}} - \frac{3}{2} \right) \right] + \langle (\delta C_s)_2 \rangle_{\mathbf{R}} \end{aligned} \quad (3.193)$$

where  $1/L_{n_s} = -(1/n_{s,0}) dn_{s,0}/dr$  and  $1/L_{T_s} = -(1/T_{s,0}) dT_{s,0}/dr$  are the equilibrium density gradient and temperature gradient scale lengths for species  $s$ , and the equilibrium drift term using  $\mathbf{v}_D$  is given by equation D.108. This is the local nonlinear gyrokinetic equation, as would be implemented in local gyrokinetic codes. The linear local gyrokinetic equation is

$$\begin{aligned} \frac{\partial h_s}{\partial t} + \frac{U_{0,\parallel}}{\sqrt{g_{\theta\theta}}} \frac{\partial h_s}{\partial \theta} + \mathbf{v}_D \cdot \frac{\partial h_s}{\partial \mathbf{R}} = \frac{q_s F_{s,0}}{T_{s,0}} \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial t} \\ - \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial y} \frac{F_{s,0}}{B_{\text{unit}}} \left[ \frac{1}{L_{n_s}} + \frac{1}{L_{T_s}} \left( \frac{\varepsilon}{T_{s,0}} - \frac{3}{2} \right) \right] + \langle (\delta C_s)_2 \rangle_{\mathbf{R}}. \end{aligned} \quad (3.194)$$

<sup>23</sup>Note the repetition of notation with Cartesian coordinates here as a consequence of convention within the literature. The context of its use should make clear which coordinate system is being referred to.

An additional advantage of the Fourier representation of fluctuating quantities in the local limit is the ability to analytically calculate gyroaverages. For both types of gyroaverage, one at constant  $\mathbf{R}$  and one at constant  $\mathbf{r}$ , this amounts to multiplying the Fourier amplitudes of the fluctuating quantity by a Bessel function  $J_0$  with an argument dependent on the magnitude of the lowest-order perpendicular wavenumber  $|\mathbf{k}_\perp| = \sqrt{g^{xx}k_x^2 + 2g^{xy}k_xk_y + g^{yy}k_y^2}$  [96, 97]. This is demonstrated explicitly in appendix D.5.3.

### 3.5.1 Normalisations

When implemented into codes, the gyrokinetic system is described using normalised quantities. The normalisations used in this work correspond to those of the code CGYRO, and are the deuterium mass  $m_D$ , the electron temperature  $T_e$ , the electron density  $n_e$ , the tokamak minor radius  $a$  and the effective magnetic field  $B_{\text{unit}}$ , defined by equation 3.190. The reference gyroradius is therefore  $\rho_{\text{unit}} = \sqrt{m_D T_e}/e B_{\text{unit}}$ , with  $\rho_* = \rho_{\text{unit}}/a$ , and the deuterium sound speed is  $c_s = \sqrt{T_e/m_D}$ . Other normalisations, useful for upcoming discussions of turbulent fluxes, include those of frequency  $c_s/a$ , particle flux  $\Gamma_{\text{GBD}} = \rho_*^2 n_e c_s$ , energy flux  $Q_{\text{GBD}} = \rho_*^2 n_e T_e c_s$  and the electrostatic potential  $\phi_{\text{unit}} = \rho_* T_e/e = B_{\text{unit}} \rho_{\text{unit}}^2 (c_s/a)$ . The normalisation of terms involving  $\mu_0$  is achieved with the unit plasma beta,  $\beta_{e,\text{unit}} = 2\mu_0 n_e T_e/B_{\text{unit}}^2$ , which is typically small in large aspect ratio tokamaks.

## 3.6 Turbulent transport in local gyrokinetics

### 3.6.1 Flux calculation

The local nonlinear gyrokinetic equation (equation 3.193) can be solved for the average turbulent properties of the plasma at a given flux surface  $r_0$ . This includes the calculation of local turbulent fluxes, of which we are particularly interested due to their role in the evolution of the transport equations used by integrated modelling suites. The form of these fluxes is now considered, initially focusing on heat flux.

Turbulent heat fluxes arise in the gyrokinetic system due to the interaction between  $\delta\phi$  and the pressure fluctuations of the present plasma species  $\delta p_s$  via the turbulent  $E \times B$  velocity [31, 71]. Once saturated, a statistically steady-state flux can be defined by taking both an ensemble average and a flux-surface average over these fluctuations, where the flux surface average of a function  $f$  is defined by [11]

$$\langle f \rangle_{\text{FS}}(r) = \frac{\int f \delta(r' - r) dV}{\int \delta(r' - r) dV} \quad (3.195)$$

where the volume  $V$  encloses the flux surface  $r$ . Of particular interest is the radial component of this averaged flux, due to its integral role in determining the confinement properties of the plasma. This is given by [31]

$$Q_s = \left\langle \nabla x \cdot \frac{\hat{\mathbf{b}} \times \nabla \delta \phi}{B} \delta p_s \right\rangle_{\text{Ens,FS}} \quad (3.196)$$

To progress with equation 3.196, we first consider the form of the flux surface average and ensemble average in the local limit.

### The local flux surface average

Evaluating equation 3.195 in the local field-aligned system for a surface labelled by  $x$ , using  $\delta(r' - r) = \delta(x' - (r - r_0)) = \delta(x' - x)$  with volume element  $dV = \mathcal{J}_{xy\theta} dx dy d\theta$  one finds

$$\begin{aligned} \langle f \rangle_{\text{FS}}(x, t) &= \frac{\int_{x_0}^{x_0+L_x} \int_{y_0}^{y_0+L_y} \int_{-\pi}^{\pi} f(x', y, \theta, t) \delta(x' - x) \mathcal{J}_{xy\theta}(\theta) dx' dy d\theta}{\int_{x_0}^{x_0+L_x} \int_{y_0}^{y_0+L_y} \int_{-\pi}^{\pi} \delta(x' - x) \mathcal{J}_{xy\theta}(\theta) dx' dy d\theta} \\ &= \frac{\int_{-\pi}^{\pi} \left[ \frac{1}{L_y} \int_{y_0}^{y_0+L_y} f(x, y, \theta, t) dy \right] \mathcal{J}_{xy\theta}(\theta) d\theta}{\int_{-\pi}^{\pi} \mathcal{J}_{xy\theta}(\theta) d\theta} \\ &= \langle f \rangle_{y,\theta}(x, t) \end{aligned} \quad (3.197)$$

where  $\langle \dots \rangle_y$  is the average over the  $y$  domain and  $\langle \dots \rangle_\theta$  is the parallel average, defined respectively by

$$\langle f \rangle_y = \frac{1}{L_y} \int_{y_0}^{y_0+L_y} f dy \quad (3.198)$$

$$\langle f \rangle_\theta = \frac{\int_{-\pi}^{\pi} f \mathcal{J}_{xy\theta}(\theta) d\theta}{\int_{-\pi}^{\pi} \mathcal{J}_{xy\theta}(\theta) d\theta}. \quad (3.199)$$

where the two averages commute.

### The local ensemble average

In the local limit, ensemble averaged quantities become constant over the intermediate length scale  $L$  in the radial direction (equation 3.171)<sup>24</sup>. Having a constant ensemble average allows an equivalence to be made between the ensemble average and the average taken over the intermediate length scale.

To illustrate this, consider only time dependence for a turbulent quantity  $f(t)$ . The key idea is that, because the ensemble average is a constant, then for a single instance of the experiment each data point for the fluctuations  $f(t_i)$  can be considered its *own* instance of

<sup>24</sup>A similar equation would apply for the binormal direction, however this is not explicit due to axisymmetry.

the experiment  $f^{(i)}$ , thus  $f(t_i) = f^{(i)}$ . Using this equivalence, from the definition of a time average we find

$$\begin{aligned}\langle f \rangle_t &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(t) dt \\ &\approx \lim_{N \rightarrow \infty} \frac{1}{N \Delta t} \sum_{i=1}^N f(t_i) \Delta t \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f^{(i)}\end{aligned}\tag{3.200}$$

which is of the form of equation 3.1. Shown explicitly for time dependence here, it is clear to see how this generalises to the two perpendicular coordinates,  $x$  and  $y$ . Note the limit of  $T \rightarrow \infty$  has been assumed, however this is also valid for  $T \gg a/c_s$ , the time scale of the fluctuations. Similarly, one requires  $L_x, L_y \gg \rho_i$ .

The local ensemble average thus consists of averaging over the radial domain, the binormal domain, and a timescale long compared to the fluctuations

$$\begin{aligned}\langle f \rangle_{\text{Ens}}(\theta) &= \frac{1}{T_0} \int_{t_0}^{t_0+T_0} \frac{1}{L_y} \int_{y_0}^{y_0+L_y} \frac{1}{L_x} \int_{x_0}^{x_0+L_x} f(x, y, \theta, t) dx dy dt \\ &= \langle f \rangle_{x,y,t}(\theta)\end{aligned}\tag{3.201}$$

where  $T_0 \gg a/c_s$ , and we have defined the radial and time averages

$$\langle f \rangle_x = \frac{1}{L_x} \int_{x_0}^{x_0+L_x} f dx\tag{3.202}$$

$$\langle f \rangle_t = \frac{1}{T_0} \int_{t_0}^{t_0+T_0} f dt.\tag{3.203}$$

### Turbulent flux calculation

Using equations 3.197 and 3.201, the combination of an ensemble average and a flux-surface average in the local limit therefore becomes

$$\langle f \rangle_{\text{FS,Ens}} = \langle f \rangle_{x,y,\theta,t}\tag{3.204}$$

which we now apply to the definition of the turbulent heat flux in equation 3.196. Evaluating equation 3.196 in the local field-aligned system, one obtains

$$Q_s = - \left\langle \frac{1}{B_{\text{unit}}} \delta p_s \frac{\partial \delta \phi}{\partial y} \right\rangle_{x,y,\theta,t}.\tag{3.205}$$

Inserting the Fourier representation from equation 3.191 for the potential and pressure fluctuations, equation 3.205 becomes

$$Q_s = \left\langle \sum_{k_x} \sum_{k'_x} \sum_{k_y} \sum_{k'_y} \frac{-ik'_y}{B_{\text{unit}}} \delta \hat{p}_{s,k_x,k_y}(\theta, t) \delta \hat{\phi}_{k'_x,k'_y}(\theta, t) e^{i(k_x+k'_x)x} e^{i(k_y+k'_y)y} \right\rangle_{x,y,\theta,t}.\tag{3.206}$$

Upon carrying out the averages in  $x$  and  $y$ , the orthogonality of the exponentials leaves only the terms with  $k'_x = -k_x$ ,  $k'_y = -k_y$ :

$$Q_s = \sum_{k_x} \sum_{k_y} \frac{ik_y}{B_{\text{unit}}} \left\langle \delta \hat{p}_{s,k_x,k_y}(\theta, t) \delta \hat{\phi}_{-k_x,-k_y}(\theta, t) \right\rangle_{\theta,t}. \quad (3.207)$$

Because the quantities involved in this calculation are real, their Fourier components satisfy  $\delta \hat{f}_{-k_x,-k_y} = \delta \hat{f}_{k_x,k_y}^*$ , which when applied to the potentials gives

$$Q_s = \sum_{k_x} \sum_{k_y} \frac{ik_y}{B_{\text{unit}}} \left\langle \delta \hat{p}_{s,k_x,k_y}(\theta, t) \delta \hat{\phi}_{k_x,k_y}^*(\theta, t) \right\rangle_{\theta,t} = \sum_{k_x} \sum_{k_y} V_{k_x,k_y} \quad (3.208)$$

where  $V_{k_x,k_y}$  has been introduced to reduce clutter in the following. Observing  $V_{k_x,k_y=0} = 0$ , equation 3.208 may be written

$$Q_s = \sum_{k_y>0} \sum_{k_x} (V_{k_x,k_y} + V_{k_x,-k_y}). \quad (3.209)$$

Using the result  $\sum_{k_x} V_{k_x,-k_y} = \sum_{k_x} V_{-k_x,-k_y}$ , as well as the property  $V_{k_x,k_y}^* = V_{-k_x,-k_y}$ :

$$Q_s = \sum_{k_y>0} \sum_{k_x} (V_{k_x,k_y} + V_{-k_x,-k_y}) = \sum_{k_y>0} \sum_{k_x} (V_{k_x,k_y} + V_{k_x,k_y}^*). \quad (3.210)$$

This may be expressed more succinctly using the definition of the real part of a complex number,  $\text{Re}[z] = (z + z^*)/2$ :

$$Q_s = 2 \sum_{k_y>0} \sum_{k_x} \text{Re}[V_{k_x,k_y}]. \quad (3.211)$$

Here the real property of the fields has introduced a symmetry to the terms, such that the flux contribution from the negative  $k_y$  modes is equal to that of the positive. Re-writing equation 3.211 using  $\text{Re}[iz] = -\text{Im}[z] = \text{Im}[z^*]$  gives

$$Q_s = 2 \sum_{k_y>0} \sum_{k_x} \frac{k_y}{B_{\text{unit}}} \left\langle \text{Im} \left[ \delta \hat{p}_{s,k_x,k_y}^*(\theta, t) \delta \hat{\phi}_{k_x,k_y}(\theta, t) \right] \right\rangle_{\theta,t}. \quad (3.212)$$

Note that the terms in this sum are related to the phase difference between the Fourier modes of the pressure and potential fluctuations by the logic presented in appendix A.3. Having now derived the expression for the turbulent transport of energy, the transport of particles<sup>25</sup> is obtained simply by replacing the pressure fluctuations with density fluctuations in equation 3.205, such that we get

$$\Gamma_s = 2 \sum_{k_y>0} \sum_{k_x} \frac{k_y}{B_{\text{unit}}} \left\langle \text{Im} \left[ \delta \hat{n}_{s,k_x,k_y}^*(\theta, t) \delta \hat{\phi}_{k_x,k_y}(\theta, t) \right] \right\rangle_{\theta,t}. \quad (3.213)$$

---

<sup>25</sup>Momentum transport is not explicitly considered in this work as due to the symmetry of the equilibria considered no momentum transport is produced.

## Flux as a spectrum of weights and saturated potentials

As shall be seen in the ensuing chapter on quasilinear models, it can be useful to consider each term in the flux sum as a product of the phase difference between the fluctuations and the magnitude of the potentials. To do so, one can define  $Z_{s,k_x,k_y}$  such that for every Fourier mode  $\delta\hat{p}_{s,k_x,k_y} = Z_{s,k_x,k_y}\delta\hat{\phi}_{k_x,k_y}$ . Multiplying both sides of this relation by  $\delta\hat{\phi}_{k_x,k_y}^*$ , taking the complex conjugate and isolating the imaginary part yields  $\text{Im} \left[ \delta\hat{p}_{s,k_x,k_y}^* (\theta, t) \delta\hat{\phi}_{k_x,k_y} (\theta, t) \right] = \text{Im} \left[ Z_{s,k_x,k_y}^* \right] \left| \delta\hat{\phi}_{k_x,k_y} \right|^2$ . The heat flux can therefore be written

$$\begin{aligned} Q_s &= \sum_{k_y > 0} 2 \sum_{k_x} \left\langle \frac{k_y \text{Im} \left[ Z_{s,k_x,k_y}^* \right]}{B_{\text{unit}}} \left| \delta\hat{\phi}_{k_x,k_y} \right|^2 \right\rangle_{\theta,t} \\ &= \sum_{k_y > 0} Q_{s,k_y}. \end{aligned} \quad (3.214)$$

Here the total flux  $Q_s$  has been decomposed into a sum of discrete flux components per positive binormal wavenumber,  $Q_{s,k_y} = 2 \sum_{k_x} \left\langle \frac{k_y \text{Im} \left[ Z_{s,k_x,k_y}^* \right]}{B_{\text{unit}}} \left| \delta\hat{\phi}_{k_x,k_y} \right|^2 \right\rangle_{\theta,t}$ . Note the appearance of the squared magnitude of the potential fluctuations  $\left| \delta\hat{\phi}_{k_x,k_y} \right|^2$ , and  $k_y \text{Im} \left[ Z_{s,k_x,k_y}^* \right] / B_{\text{unit}}$ , which contains the description of the phase difference between the fluctuations.

The flux calculation can then be expressed as a 1D sum over  $k_y$  in terms of average phases and potential magnitudes. Here we multiply and divide each term in the 1D flux sum of equation 3.214 by the magnitude of the 1D potentials at that  $k_y$ , which from Parseval's theorem (appendix E.1) is

$$\left\langle \left| \delta\hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t} = \sum_{k_x} \left\langle \left| \delta\hat{\phi}_{k_x,k_y} \right|^2 \right\rangle_{\theta,t}. \quad (3.215)$$

Hence the flux calculation becomes

$$Q_s = 2 \sum_{k_y > 0} W_{s,k_y} \left\langle \left| \delta\hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t} \quad (3.216)$$

where we have defined the ‘weight’  $W_{s,k_y}$ , such that

$$W_{s,k_y} = \frac{\sum_{k_x} \frac{k_y}{B_{\text{unit}}} \left\langle \text{Im} \left[ Z_{s,k_x,k_y}^* \right] \left| \delta\hat{\phi}_{k_x,k_y} \right|^2 \right\rangle_{\theta,t}}{\sum_{k_x} \left\langle \left| \delta\hat{\phi}_{k_x,k_y} \right|^2 \right\rangle_{\theta,t}} = \frac{\frac{1}{2} Q_{s,k_y}}{\left\langle \left| \delta\hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \quad (3.217)$$

named as such due to it ‘weighting’ the contributions of the magnitude of the potential fluctuations to the total flux. Note the factor of  $\frac{1}{2}$  in the final expression of the weights due to historical convention. With reference to equation 3.216, we see that the phase relation between the fluctuations is described entirely by the weights.

### 3.6.2 Turbulent flux characteristics

#### Zero non-adiabatic transport

It was previously shown that the first order fluctuating distribution for a given species  $\delta f_{s,1}$  contains an adiabatic part  $\propto \delta\phi_1$  and a non-adiabatic part,  $h_s$  (equation 3.96). Thus each of the velocity moments of  $\delta f_{s,1}$  will contain an adiabatic part and a non-adiabatic part. For example, the density fluctuations are of the form  $\delta n_{s,1} = -(q_s n_{s,0}/T_{s,0}) \delta\phi_1 + \int \langle h_s \rangle_{\mathbf{r}} d^3\mathbf{v}$ , with similar results for the momentum and pressure fluctuations.

Inserting velocity moment fluctuations proportional to  $\delta\phi_1$  into the definitions of the turbulent fluxes, such as equation 3.213, one has terms of the form  $\propto \left\langle \text{Im} \left[ \delta\hat{\phi}_{k_x, k_y}^* \delta\hat{\phi}_{k_x, k_y} \right] \right\rangle_{\theta, t} \propto \left\langle \text{Im} \left[ \left| \delta\hat{\phi}_{k_x, k_y} \right|^2 \right] \right\rangle_{\theta, t} = 0$ . We see therefore that the velocity moment fluctuations associated with the adiabatic part of the distribution do not produce any transport [31], as there is no imaginary part to the phase difference with the potential<sup>26</sup>. All turbulent transport in gyrokinetics is therefore associated with the non-adiabatic part of the fluctuating distribution,  $h_s$ .

#### Ambipolar particle transport

The turbulent transport of particles in gyrokinetics is intrinsically ambipolar [98], the quality by which the plasma species move so as to retain zero net charge density, as a result of quasineutrality. Using  $\sum_i Z_i \delta n_i = \delta n_e$  from equation 3.104 and the particle flux analogue of equation 3.196, we find

$$\begin{aligned} \Gamma_e &= \left\langle \nabla x \cdot \frac{\hat{\mathbf{b}} \times \nabla \delta\phi}{B} \delta n_e \right\rangle_{\text{Ens,FS}} \\ &= \sum_i Z_i \left\langle \nabla x \cdot \frac{\hat{\mathbf{b}} \times \nabla \delta\phi}{B} \delta n_i \right\rangle_{\text{Ens,FS}} \\ &= \sum_i Z_i \Gamma_i \end{aligned} \quad (3.218)$$

and so the particle transport of the electrons is balanced by the charge-weighted particle transport of the ions, where the transport is a result of the non-adiabatic part of the distribution<sup>27</sup>.

In the case of a single hydrogenic ion species, this condition simply reduces to  $\Gamma_e = \Gamma_i$ .

Note the ambipolarity constraint applies only to the particle transport, and not to the higher velocity moments of momentum and energy.

<sup>26</sup>Note the similarity of this result with the electron drift wave analysis in section 2.3.1, in which it was found no instability was incurred when the electrons were modelled with an adiabatic response.

<sup>27</sup>Thus for a simulation using adiabatic electrons, one has  $\sum_i Z_i \Gamma_i = 0$ .

## 3.7 Outputs of the local gyrokinetic system

### 3.7.1 Nonlinear gyrokinetics

#### Flux spectra

Some graphical examples of outputs from local gyrokinetic simulations will now be presented. From nonlinear gyrokinetics, figure 3.6a shows a time trace of the total ion heat flux  $Q_i$  for the ‘GA standard’ (GA-std) case [99] in deuterium using CGYRO, equal to equation 3.214 *before* the averaging over time is applied. At the start of the simulation, the ‘linear’ regime of the turbulence can be observed, in which the flux and the associated fluctuations grow exponentially, before entering the nonlinear regime characterised by steady levels of fluctuations, indicating turbulent saturation.

Also exhibited in figure 3.6a is the methodology of time averaging used in this work. Nonlinear simulations have been run to a time  $T \gg a/c_s$ , typically  $\sim 1000$ . In order to have a representative measurement and error, three averages of the trace are taken over three equally-sized windows of  $0.28T$  between  $0.16T$  and  $T$ . This avoids contributions from the linear regime. The mean of these three measurements is then taken for the final value, with a standard deviation calculated from the values of the three windows<sup>28</sup>. Any error bars displayed in this work are calculated via this method unless stated otherwise. It can be shown that the mean of the three measurements is equivalent to taking a single measurement between  $0.16T$  to  $T$ , such as in equation 3.203, via

$$\frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} f dt = \frac{1}{N} \sum_{j=1}^N \frac{1}{(\Delta t/N)} \int_{t_0+(j-1)(\Delta t/N)}^{t_0+j(\Delta t/N)} f dt \quad (3.219)$$

for any integer  $N$ .

The total flux  $Q_s$  can be broken down into a spectrum of 1D flux components  $Q_{i,k_y}$ , as per equation 3.214. Figure 3.6b exhibits the spectrum of flux components for the ion heat flux of the GA-std case, in which a peak is seen in the region  $k_y \rho_i \sim 0.2$ , characteristic of ion-scale turbulence. Summing each flux component in the spectrum equates to the final value from figure 3.6a.

---

<sup>28</sup>A compromise must be made between the number of averaging windows and the extent of the windows necessary to satisfy  $\gg a/c_s$  in order to obtain a representative value of the uncertainty in the mean of the fluctuations.

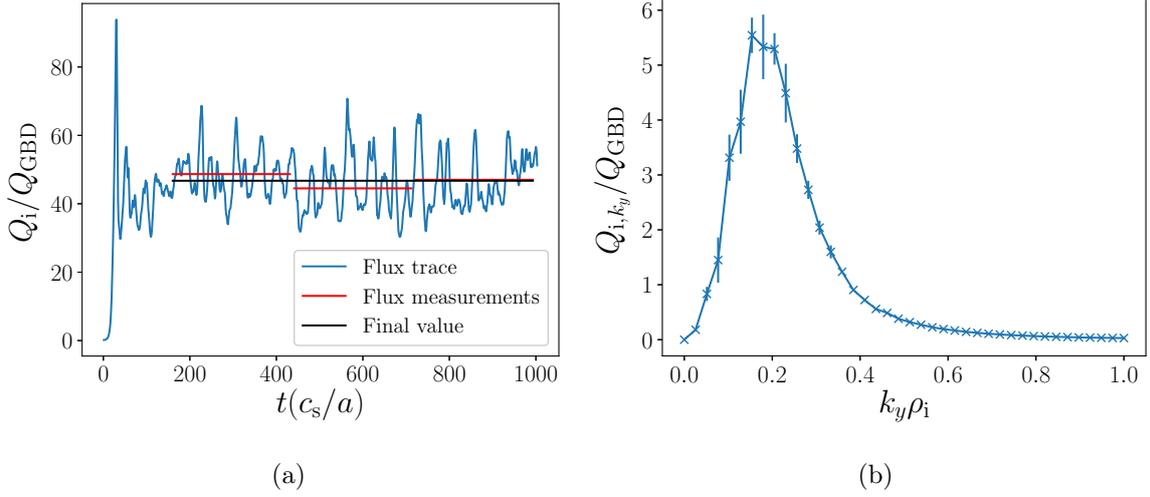


Figure 3.6: (a) The trace of the total ion heat flux in the GA-std case in D, exhibiting the method of time averaging used in this work. (b) The flux component spectrum  $Q_{i,k_y}$  of the ion heat flux in D for the GA-std case.

### Potential spectra

Shown in figure 3.7 is the spectrum of saturated potential magnitudes for the same GA-std simulation. Both subfigures plot the same spectrum, however in figure 3.7b the  $k_y = 0$  mode is not present so as to more clearly demonstrate the spectrum for  $k_y > 0$ . The  $k_y = 0$  mode is known as the *zonal mode*, and as can be seen in equation 3.214, never contributes to transport, however mechanisms have been presented whereby the zonal mode plays a role in the dynamics of turbulent saturation [100, 101]. Turbulent transport is therefore associated with the potential spectrum for  $k_y > 0$ . Note again the peak in the spectrum around  $k_y \rho_i \sim 0.2$ , reminiscent of the flux components in figure 3.6b.

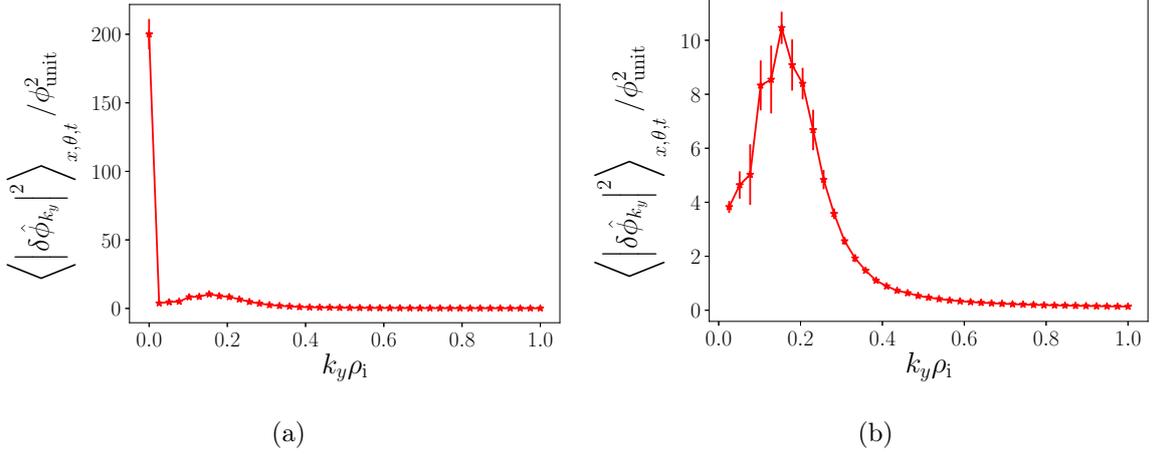


Figure 3.7: The 1D saturated potential spectrum for the GA-std case in D, shown with (a) and without (b) the zonal mode.

### Weight spectrum

Using the flux component spectrum of figure 3.6b and the saturated potential spectrum from figure 3.7b one can calculate the weights for the GA-std simulation, using equation 3.217. As these are calculated from nonlinear gyrokinetics, we dub these the ‘nonlinear weights’  $W_{i,k_y}^{\text{NL}}$ , in anticipation of future generalisation. These are shown in figure 3.8.

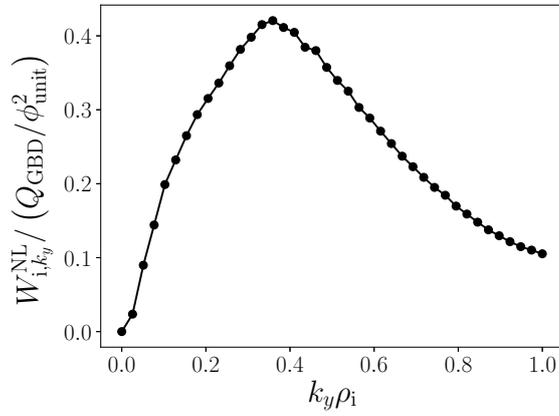


Figure 3.8: Nonlinear weight spectrum for the GA-std case in D.

### 3.7.2 Linear gyrokinetics

#### Eigenvalues

Turning now to the linear gyrokinetic system, the absence of the nonlinear term causes there to be no saturation in the system<sup>29</sup>. Typically using an explicit time integration scheme, the fluctuating quantities simply continue to oscillate at the real part of the frequency  $\tilde{\omega}_{k_y}$ , and continue to grow to arbitrary size via the growth rate  $\gamma_{k_y}$ , which are the eigenvalues that characterise the dominant mode<sup>30</sup> of the instability. By conducting multiple linear simulations across a range of  $k_y$  values, one can generate frequency spectra of these eigenvalues, shown in figure 3.9. In figure 3.9a the growth rate spectrum for the GA-std case is shown, with the corresponding real frequencies shown in figure 3.9b. The sign of  $\tilde{\omega}_{k_y}$  can be used to probe the turbulence mode present in the simulation. Here the negative frequency represents an ion-dominated mode, corresponding to ITG turbulence indicative of the GA-std case<sup>31</sup>.

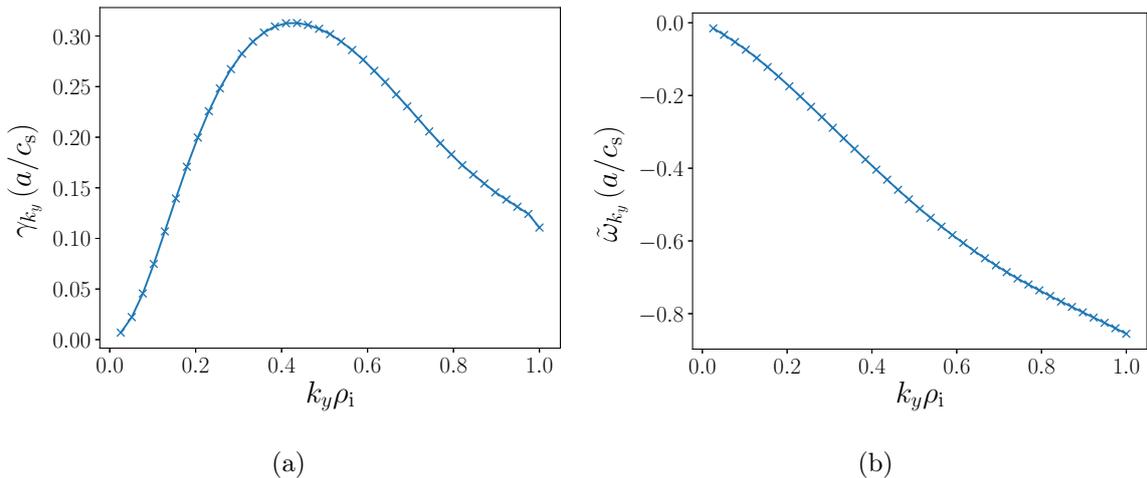


Figure 3.9: Eigenvalue spectra for the linear gyrokinetic system of the GA-std case in D, showing the growth rates  $\gamma_{k_y}$  (a) and the real frequencies  $\tilde{\omega}_{k_y}$  (b).

#### Eigenfunctions and linear weights

Because there is no turbulent saturation in a linear simulation, the eigenfunctions, such as the potential perturbations  $\delta\phi_{k_x, k_y}(\theta, t)$ , become arbitrary large due to their growth rates  $\propto e^{\gamma_{k_y} t}$ . However as all quantities grow with the same growth rate, this does mean that physically-

<sup>29</sup>This also allows for linear simulations to be conducted at a single value of  $k_y$ .

<sup>30</sup>The mode with the largest growth rate for a given  $k_y$ .

<sup>31</sup>Note some codes can use the opposite convention, in which a positive frequency indicates an ion mode and a negative frequency indicates an electron mode.

meaningful *ratios* of eigenfunction quantities may be taken to gain further information about the system.

Consider the definition of the heat flux from equation 3.212. In the linear system this becomes arbitrary large,  $Q_{s,k_y} \propto \left\langle \text{Im} \left[ \delta \hat{p}_{s,k_x,k_y}^* (\theta, t) \delta \hat{\phi}_{k_x,k_y} (\theta, t) \right] \right\rangle_{\theta,t} \propto e^{2\gamma_{k_y} t}$ . If we were to divide this quantity by  $\langle |\delta \hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t} \propto e^{2\gamma_{k_y} t}$ , the exponential growth would cancel, leaving a defined value for the ‘linear flux per unit potential magnitude squared’,  $Q_{s,k_y} / \langle |\delta \hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t}$ . From equation 3.217, we know such a quantity to be the definition of the weight, which is related to the phase difference between the fluctuations. We call this the ‘linear weight’,  $W_{s,k_y}^L$ . The linear weights calculated from linear simulations of the GA-std case are shown in figure 3.10. Note the similarity in the shape of the spectrum between the linear weights and the nonlinear

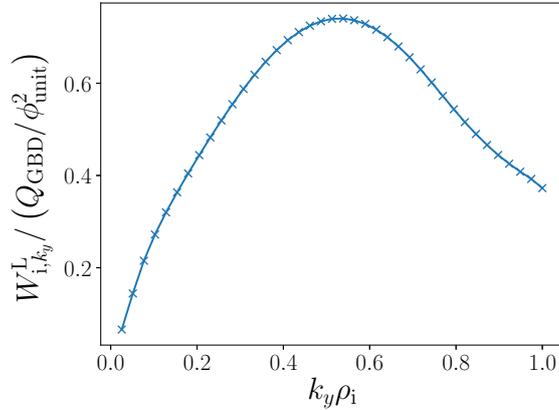


Figure 3.10: The linear weight spectrum for the ion heat flux of the GA-std case in D.

weights (figure 3.8). It is this similarity that forms the basis of quasilinear turbulence models, which are the focus of the next chapter.

# Chapter 4

## Quasilinear models

### 4.1 Introduction

In section 2.5.1 the use case was presented for quasilinear turbulence models, with them bridging a compromise between the need for the calculation of local turbulent fluxes in integrated modelling and the timescale on which they are required, typically  $\mathcal{O}(1\text{ s})$ . This reduction in computational resources relative to nonlinear gyrokinetics is achieved by the use of several simplifications assumed in quasilinear models. Having considered the general case of the calculation of local nonlinear gyrokinetic fluxes in section 3.6, the mathematical structure by which quasilinear models achieve this reduction will now be demonstrated. The specifics of two state of the art quasilinear models will then be detailed, TGLF [50] and QuaLiKiz [49], after which a comparison will be shown between the isotope scaling of the fluxes of these models and those from nonlinear gyrokinetics. The discrepancies between the models types will then be analysed to determine their physical origin.

### 4.2 Quasilinear theory

To describe the mathematical structure of quasilinear models we first consider the local heat flux<sup>1</sup>. In local nonlinear gyrokinetics it is calculated by equation 3.216, repeated here for convenience with the ‘NL’ label for the weights, following the discussion of the nonlinear

---

<sup>1</sup>The following discussion also relates to the particle and momentum flux by replacing instances of ‘pressure fluctuations’ with density and momentum fluctuations respectively.

weights and linear weights in section 3.7:

$$\begin{aligned}
Q_s &= \sum_{k_y > 0} Q_{s,k_y} \\
&= 2 \sum_{k_y > 0} W_{s,k_y}^{\text{NL}} \left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}.
\end{aligned} \tag{4.1}$$

This flux calculation comprises a 1D sum over  $k_y$  of the product of two spectra: the nonlinear weights  $W_{s,k_y}^{\text{NL}}$  and the saturated potential  $\langle |\delta \hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t}$ . Quasilinear models are required to approximate both of these spectra for a given local equilibrium without solving the local nonlinear gyrokinetic system. The calculation of the weight spectrum will first be considered.

#### 4.2.1 The weight spectrum, $W_{s,k_y}$

##### The quasilinear approximation

The nonlinear weights represent the average phase relation between the potential and pressure fluctuations during the nonlinear phase of the turbulence. As discussed in section 3.7.2, a similar quantity can be defined for the local linear gyrokinetic system, the linear weights  $W_{s,k_y}^{\text{L}}$ . Provided that the linear weights and nonlinear weights roughly correlate, these can be used to provide an approximation of the nonlinear weight spectrum. This illustrates the central requirement of the quasilinear theory, which is the so-called *quasilinear approximation*: the average phase between the fluctuations in the linear regime of the turbulence is approximately preserved in the nonlinear phase, such that

$$W_{s,k_y}^{\text{L}} \approx W_{s,k_y}^{\text{NL}}. \tag{4.2}$$

Note that this assumption applies for all species, at all values of  $k_y$ , and for all velocity moments. Previous studies validating the quasilinear approximation have predominantly shown it to be a reasonable assumption for core turbulence [51, 102, 103].

##### Reduced linear equations

Solving the linear gyrokinetic system to obtain the weight spectrum still remains too costly for the volumes of calculations required for routine use in integrated modelling, and so quasilinear models will typically make further reductions to the linear gyrokinetic system to speed up calculations. The philosophy of these reductions must be that they produce sufficient speed-up in the solution for the linear properties of the turbulence, whilst remaining accurate in comparison to the ‘correct’ linear gyrokinetic result. Such reductions can include making

assumptions on the form of the eigenstate, or solving gyrofluid equations<sup>2</sup> [104, 105] subject to appropriate closure instead of the gyrokinetic system. Introducing the superscript ‘RL’ (reduced linear) to denote this reduced linear system, in order to maintain a reasonable approximation of the weight spectrum one requires from these reductions

$$W_{s,k_y}^{\text{RL}} \approx W_{s,k_y}^{\text{L}}. \quad (4.3)$$

Combined with the quasilinear approximation this should allow the reduced linear weights, which are able to be calculated on a timescale commensurate with integrated modelling requirements, to be similar to those of the nonlinear weights  $W_{s,k_y}^{\text{RL}} \approx W_{s,k_y}^{\text{NL}}$ .

### Eigenvalues

In solving for the linear dynamics of the present microinstabilities one not only obtains the linear weights but also the linear eigenvalues, namely the real frequency  $\tilde{\omega}_{k_y}$  and the linear growth rate  $\gamma_{k_y}$  of the mode. Whilst these eigenvalues are not explicitly present in the calculation of equation 4.1, these can help to inform on the nature of the underlying instability, and are commonly used in the quasilinear calculation of  $\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t}$ . Note that for an appropriate reduced linear system an analogous condition of equation 4.3 also applies to the eigenvalues,  $\gamma_{k_y}^{\text{RL}} \approx \gamma_{k_y}^{\text{L}}$  and  $\tilde{\omega}_{k_y}^{\text{RL}} \approx \tilde{\omega}_{k_y}^{\text{L}}$ .

### Subdominant modes

An upshot of the reduced linear system is that the smaller dimensionality employed typically makes it easier to solve for eigenvalues directly, rather than finding a solution via the fastest growing mode with an initial value solver. Thus quasilinear models can generally more readily obtain information on subdominant modes than linear gyrokinetic simulations, and include their contribution in the calculation of turbulent fluxes [106, 107].

These subdominant modes can be labelled with the index  $m$ , where  $m = 1$  corresponds to the fastest growing eigenmode (the dominant mode),  $m = 2$  the second fastest, and so on, such that growth rates for a given  $k_y$  are labelled  $\gamma_{k_y,m}$ , and reduced linear weights  $W_{s,k_y,m}^{\text{RL}}$ . These contributions are considered in section 4.3.1.

---

<sup>2</sup>These are analogous to the fluid equations discussed in section 1.2.1, however they are obtained by taking velocity moments over the gyrokinetic equation rather than the kinetic equation.

## 4.2.2 The saturated potential spectrum, $\langle |\delta \hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t}$

### Quasilinear saturation rules

The process by which turbulence saturates is inherently nonlinear, and thus an approximation of the saturated potential spectrum cannot be obtained from solutions to the linear system. This is the purpose of a *saturation rule*, a semi-empirical function that describes the magnitude and shape of the saturated potential spectrum in quasilinear models. These functions are typically informed by a combination of analytic results, physical approximations and nonlinear gyrokinetic simulations.

Figure 3.7b shows an example of a saturated potential spectrum obtained from nonlinear gyrokinetics for the GA-std case in D, for  $k_y > 0$ . Having remarked how the zonal mode ( $k_y = 0$ ) never produces transport, this is not something that is essential to model for the calculation of turbulent fluxes<sup>3</sup>. From the figure we note its initial rise to a peak, characteristic of saturated potential spectra, before tailing off for larger values of  $k_y$ . We split the considerations of a saturation rule into the ‘spectral shape’, which describes how the function varies with  $k_y$ , and the ‘saturation level’, which characterises the ‘height’ of the spectrum. The saturation rule can be informed by the linear analysis through taking the eigenvalues as arguments.

Of course saturation rules must not just recreate the GA-std in D spectrum, but the spectrum of any equilibrium being considered. For a parameter space as broad as that of local tokamak equilibria (including different plasma species, Maxwellian gradients, general geometry, at both ion and electron scales) this seems a daunting task. The vastness of the parameter space coupled with the expense of nonlinear gyrokinetic simulations encourages saturation rules to be built on physical principles rather than being primarily fitted to databases, as these generally have more predictive power and greater ability in extrapolation. Such principles are called *saturation mechanisms*, and describe the physical processes by which turbulence saturates in the nonlinear system. A rudimentary saturation mechanism example will now be considered.

### The mixing length approximation

A common ingredient used in saturation rules historically is the mixing length approximation [25, 49], which uses arguments regarding the correlation length and time scales of the turbulence to predict the magnitude of the peak of the saturated potential spectrum. Modelling

---

<sup>3</sup>Some models may still choose to include the calculation of the zonal potential due to assumptions regarding its influence on turbulent saturation through ‘zonal flows’. An example of this is discussed in section 4.4.1.

the transport as a diffusive process one has for particle transport  $\Gamma \approx D_{\text{turb}} n_0 / a$ . Assuming a dominant contribution from the peak of the potential spectrum and assuming a small phase shift between the density and potential fluctuations<sup>4</sup> for equation 3.213, one obtains (suppressing the angular averaging brackets)

$$D_{\text{turb}} \sim \frac{a\Gamma}{n_0} \propto \left| \delta \hat{\phi}_{k_y} \right|_{\text{max}}^2 \quad (4.4)$$

where  $\left| \delta \phi_{k_y} \right|_{\text{max}}^2$  is the peak of the potential spectrum. Expressing  $D_{\text{turb}}$  in terms of an eddy turnover time  $\sim 1/\gamma_{k_y}$  and step size  $\sim 1/k_{\perp}$  via a random walk argument, we have

$$\left| \delta \phi_{k_y} \right|_{\text{max}}^2 \propto \frac{\gamma}{k_{\perp}^2} \Big|_{\text{max}}. \quad (4.5)$$

This is the mixing length approximation, from which we can see an example of how the form of the nonlinear spectrum can be informed by the linear dynamics of the instability. Having specified the potential magnitudes at the spectrum peak, the form of the spectrum with  $k_y$  can then be constructed around it.

### $E \times B$ shear

In this work so far no equilibrium electric fields have been assumed,  $\mathbf{E} = \mathbf{0}$ . Their inclusion generates equilibrium  $E \times B$  shear, a physical mechanism associated with plasma rotation by which turbulent eddies can become decorrelated and turbulent transport suppressed [108, 109], as well as cause the necessary symmetry-breaking to generate momentum transport [110–112]. The inclusion of  $E \times B$  shear in quasilinear models is typically parameterised through an equilibrium input parameter, the shearing rate  $\gamma_E$ , and the consequences of non-zero  $E \times B$  shear must be explicitly defined in the saturation rule.

## 4.3 Fluxes in reduced models

We consider two types of reduced models in this work: those with ‘exact’ linear solvers, which compute the linear behaviour of microinstabilities using the linear gyrokinetic system (L), and those with ‘fast’ linear solvers, which solve the reduced linear system (RL). The results of these solvers are then combined with a saturation rule in order to calculate turbulent fluxes. While exact linear solvers are too costly for use in integrated modelling they are very useful for the purposes of diagnosing issues with saturation rules and thus saturation rule development, as the dynamics of the linear microinstabilities are computed more accurately. In comparing

---

<sup>4</sup>For example the form of equation 2.10 using  $\Delta Z = 1 - i|\delta|$ ,  $|\delta| \ll 1$ .

the results of a saturation rule paired with an exact linear solver to nonlinear gyrokinetic data, one knows that any discrepancies must result from the form of the saturation rule, not from errors associated with the linear arguments. Models of this type have been used previously [113–116]. We therefore define a quasilinear saturation rule using the full linear gyrokinetic inputs, but it is often evaluated in integrated modelling using the reduced linear eigenvalues.

### 4.3.1 Flux calculation

The flux calculation of quasilinear models using an exact linear solver will first be considered. Starting from equation 4.1 and initially assuming only the dominant mode is considered in the linear solver, one multiplies and divides by the linear weight to get

$$\begin{aligned} Q_s &= 2 \sum_{k_y > 0} \left[ \frac{W_{s,k_y}^{\text{NL}}}{W_{s,k_y}^{\text{L}}} \right] W_{s,k_y}^{\text{L}} \left\langle |\delta\hat{\phi}_{k_y}|^2 \right\rangle_{x,\theta,t} \\ &= 2 \sum_{k_y > 0} \Lambda_{s,k_y} W_{s,k_y}^{\text{L}} \left\langle |\delta\hat{\phi}_{k_y}|^2 \right\rangle_{x,\theta,t} \end{aligned} \quad (4.6)$$

where  $\Lambda_{s,k_y} = W_{s,k_y}^{\text{NL}}/W_{s,k_y}^{\text{L}}$  is the ‘quasilinear approximation (QLA) function’. This defines an explicit measure of the quasilinear approximation, equation 4.2, is typically assumed to be 1 in current quasilinear models.

For the purposes of numerical modelling, one does not mathematically describe the potentials  $\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t}$ , but rather the potentials normalised to the binormal grid spacing  $\Delta k_y$ , as this can be shown to be invariant under a change of grid resolution once past the point of convergence (Appendix E.2). This consideration yields

$$Q_s = 2 \sum_{k_y > 0} \Lambda_{s,k_y} W_{s,k_y}^{\text{L}} \left[ \frac{\left\langle |\delta\hat{\phi}_{k_y}|^2 \right\rangle_{x,\theta,t}}{\Delta k_y} \right] \Delta k_y \quad (4.7)$$

which is the final result for the structure of fluxes relevant to quasilinear models with an exact linear solver. Each term in the sum over  $k_y$  is comprised of four parts, two of which are obtained from linear simulations, and two of which must be prescribed. The two quantities directly obtained from the linear physics are the adopted binormal grid spacing  $\Delta k_y$  and the linear weight,  $W_{s,k_y}^{\text{L}}$ . The two quantities in need of prescription are the grid-independent potentials  $\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t}/\Delta k_y$ , calculated by the quasilinear saturation rule, and the QLA function,  $\Lambda_{s,k_y}$ .

The flux calculation for the fast model type is similar to equation 4.7, however as they solve a reduced linear system, some additional sources of error are incurred due to their approximate

calculation of the linear response. These errors can enter through the linear weights  $W_{s,k_y}^L \rightarrow W_{s,k_y}^{\text{RL}}$  and through linear quantities used as arguments in the saturation rule,  $\gamma_{k_y}^L \rightarrow \gamma_{k_y}^{\text{RL}}$ . Including contributions from the subdominant modes  $m$ , equation 4.7 generalises to

$$Q_s = 2 \sum_m \sum_{k_y > 0} \Lambda_{s,k_y,m} W_{s,k_y,m}^{\text{RL}} \left[ \frac{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}}{\Delta k_y} \right]_m \Delta k_y. \quad (4.8)$$

Having analysed the general formulation of quasilinear models, the specifics of two state-of-the-art models, TGLF and QuaLiKiz, will now be discussed.

## 4.4 TGLF and QuaLiKiz

### 4.4.1 TGLF

TGLF (‘Trapped Gyro-Landau Fluid’) is a quasilinear turbulence model which uses gyrofluid equations for its evaluation of the linear fluctuations, developed from its predecessor GLF [117]. For this work we consider the linear solvers of quasilinear models to be a ‘black box’ from which we obtain the reduced linear weights and growth rates, and thus for details on the derivation and numerical schemes involved in these equations the reader is referred to [50]. We note however that included in TGLF is the physics of general flux-surface shaping using the Miller parameterisation, subdominant modes, plasma rotation, electromagnetic fluctuations and operation with an arbitrary number of species. Of TGLF’s three saturation rules we shall consider the latest two, SAT1 and SAT2.

#### SAT1

SAT1 [118] is based on the saturation mechanism of zonal mixing, motivated by observations made in multiscale simulations using the gyrokinetic code GYRO [119]. It is argued that between the mechanisms of zonal flow mixing and drift wave mixing, which respectively account for the influence of the  $k_y = 0$  and  $k_y \neq 0$  modes of the potential in the nonlinear term of the gyrokinetic equation, only the zonal flow mixing mechanism can compete with the linear growth of the electron-scale modes and thus forms the dominant saturation mechanism. The zonal flow mixing rate is given by  $k_y V_{\text{ZF}}$ , where  $V_{\text{ZF}}$  is the zonal flow velocity. This is determined quasilinearly by the balancing of the linear growth and the zonal flow mixing rate, such that

$$V_{\text{ZF}} = \frac{\gamma_{k_y}}{k_y} \Big|_{\text{max}} = \frac{\gamma_{\text{max}}}{k_{\text{max}}} \quad (4.9)$$

which is found to be in good agreement with nonlinear gyrokinetic results<sup>5</sup>. Here  $k_{\max}$  is defined by

$$\left. \frac{d}{dk_y} \left( \frac{\gamma_{k_y}}{k_y} \right) \right|_{k_y=k_{\max}} = 0 \quad (4.10)$$

with the corresponding growth rate

$$\gamma_{\max} = \gamma_{k_y=k_{\max}}. \quad (4.11)$$

A graphical example of the determination of these quantities for a given equilibrium is shown in figure 4.1.

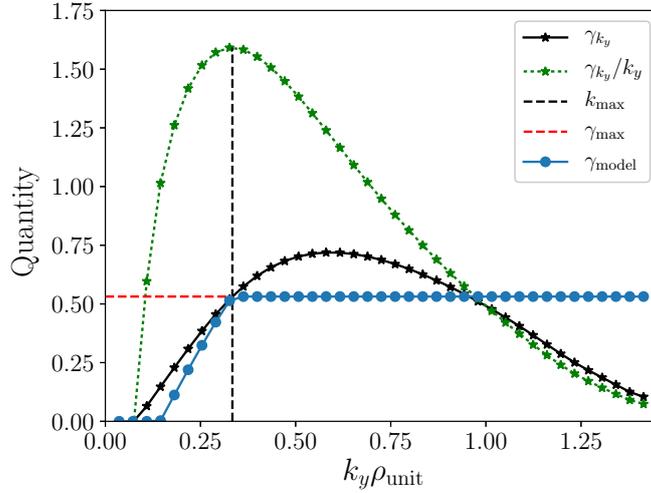


Figure 4.1: Demonstration of the calculation of  $k_{\max}$  and  $\gamma_{\max}$  from a linear growth rate spectrum, here for the  $a/L_n = 3.0$  case in H (section 4.6). Starting from  $\gamma_{k_y}$  (solid black), one divides this spectrum by  $k_y$  to obtain  $\gamma_{k_y}/k_y$  (dotted green). The maximum of this curve is then found via quadratic interpolation, with the  $k_y$  value at which this occurs being defined as  $k_{\max}$ . Returning then to  $\gamma_{k_y}$ , the value of this curve at  $k_{\max}$  is defined as  $\gamma_{\max}$ . These quantities are then used in the construction of  $\gamma_{\text{model}}$  (blue circles), given by equation 4.12.

The interplay between the linear growth of the modes and the suppression due to zonal mixing is captured by the effective growth rate  $\gamma_{\text{model}}$ ,

$$\gamma_{\text{model}} = \begin{cases} \text{MAX}(\gamma_{k_y} - 0.48V_{\text{ZF}}(k_{\max} - k_y), 0.0) & k_y < k_{\max} \\ \gamma_{\max} + \text{MAX}(\gamma_{k_y} - V_{\text{ZF}}k_y, 0.0) & k_y \geq k_{\max}. \end{cases} \quad (4.12)$$

<sup>5</sup>Typically there will be two locations at which the spectrum of  $\gamma_{k_y}/k_y$  peaks, one in the ion scale and one in the electron scale. The one in the ion scale is always used, as the  $\gamma_{\max}$  obtained from the electron scale is too large to agree with nonlinear GYRO simulations.

where the factor of 0.48 has been set via comparison to gyrokinetic simulations. Note the multiscale effect through the ion-scale saturation ( $\gamma_{\max}$ ) playing a role in the growth of the electron-scale modes.

The form of the potential spectrum is similar to that obtained from critical balance invoked between the time derivative and the nonlinear term in the gyrokinetic equation [120, 121]. Assuming  $\partial/\partial t \rightarrow \gamma$ , and  $\delta\mathbf{v}_E \cdot \nabla \sim k_x k_y \delta\phi/B$ , the balancing of these terms obtains  $\delta\phi^2 \sim B^2 \gamma^2 / (k_x k_y)^2$ . Using the effective growth rate, the observation that the root mean square radial wavenumber  $k_{x,\text{RMS}}$  approximately follows  $k_y/k_{x,\text{RMS}} = 0.56 \text{MAX} \left[ \left( \frac{k_y \rho_{\text{unit}}}{1.9}, 1.0 \right)^{0.5} \right]$ , and including a factor of  $\rho_i$  to account for the  $\Delta k_y$  normalisation, SAT1 models the potential spectrum as

$$\left. \frac{|\delta\phi_{k_y}|^2}{\Delta k_y} \right|_{\text{SAT1}} = 14.9 \rho_i \left( \frac{\gamma_{\text{model}}}{k_y^2} \text{MAX} \left[ \left( \frac{k_y \rho_{\text{unit}}}{1.9} \right)^{0.5}, 1.0 \right] \right)^2 B_{\text{unit}}^2 \quad (4.13)$$

where the prefactor is set to a database of gyrokinetic simulations and the factor of  $B_{\text{unit}}^2$  accounts for the equilibrium magnetic field dependence.

The saturated potential for subdominant mode  $m$  is given by that of the dominant mode  $m = 1$  multiplied by a weighting, based on the subdominant growth rate  $\gamma_{k_y,m}$  relative to the dominant growth rate,  $\gamma_{k_y,m=1}$ . Their ratio is raised to the fourth power to give

$$\left. |\delta\phi_{k_y,m}|^2 \right|_{\text{SAT1}} = \left. |\delta\phi_{k_y,m=1}|^2 \right|_{\text{SAT1}} \left( \frac{\gamma_{k_y,m}}{\gamma_{k_y,m=1}} \right)^4. \quad (4.14)$$

where the potential of the dominant mode is given by equation 4.13.

The effect of  $E \times B$  shear in SAT1 is based on the *spectral shift* model [122, 123], which makes advancements on the previous paradigm of a 1D ‘quench’ rule [124] by modelling the 2D saturated potentials in  $k_x$  and  $k_y$ , which is assumed to have a squared Lorentzian form. The estimation of the resulting radial shift of the spectrum is used to suppress the saturated potential spectrum as well as introduce asymmetry to the linear system, allowing for the generation of non-zero momentum transport.

## SAT2

SAT2 [52] represents an evolution of SAT1. It is still based on the saturation mechanism of zonal flow mixing, however has been extended to incorporate physical observations of turbulent spectra made since SAT1’s creation. The model for the RMS  $k_x$  has been generalised to include

flattening in the low  $k_y$  region, now modelled by

$$k_x^{\text{model}} = \begin{cases} \frac{0.76k_{\text{max}}}{\sqrt{g^{xx}(\theta=0)}} & k_y < 0.76k_{\text{max}} \\ \frac{0.76k_{\text{max}}}{\sqrt{g^{xx}(\theta=0)}} + \frac{1.22}{\sqrt{g^{xx}(\theta=0)}} \frac{B(\theta=0)}{B_{\text{unit}}} (k_y - 0.76k_{\text{max}}) & k_y \geq 0.76k_{\text{max}}. \end{cases} \quad (4.15)$$

The model for the effective growth rate is now

$$\gamma_{k_y}^{\text{eff}} = \begin{cases} b_2 \gamma_{k_y} & k_y < k_{\text{max}} \\ b_2 \gamma_{\text{max}} & k_y \geq k_{\text{max}} \end{cases} \quad (4.16)$$

for  $(b_2)^2 = 3.74 (12L_p/R_0)$ , where  $L_p$  is the length scale of the pressure gradient, obtained via fitting to a database of nonlinear simulations. SAT2 continues to assume a Lorentzian shape in radial wavenumber  $k_x$  and now captures the  $\theta$  variation in the spectrum via the shape function  $G^2(\theta)$ , for which further details can be found in [52]. Bringing this all together, including averaging over the poloidal direction and a factor of  $k_{\text{max}}$  for the  $\Delta k_y$  normalisation, the model for the saturated potentials is SAT2 is given by

$$\left. \frac{|\delta\phi_{k_y}|^2}{\Delta k_y} \right|_{\text{SAT2}} = \frac{\langle G^2(\theta) \rangle_{\theta}}{k_{\text{max}}} \left( \frac{\gamma^{\text{eff}}}{k_x^{\text{model}} k_y} \right)^2. \quad (4.17)$$

The effect of subdominant modes is carried forward from SAT1, described by equation 4.14. To present the model for the  $E \times B$  shear one defines the effect of the shear  $F_{s,k_y}$  as the ratio of flux components with shear to those without  $F_{s,k_y}(\gamma_E) = Q_{s,k_y}(\gamma_E)/Q_{s,k_y}(\gamma_E = 0)$ . The SAT2 version of the spectral shift model is

$$F_{s,k_y}(\gamma_E) = \frac{\tilde{W}_{s,k_y,k_x=k_{x0}}^L / \tilde{W}_{s,k_y,k_x=0}^L}{\left(1 + \left(\alpha_x \frac{k_{x0}}{k_x^{\text{model}}}\right)^{\sigma_x}\right)^2 \left(1 + \left(\frac{k_{x0}}{k_x^{\text{model}}}\right)^2\right)}. \quad (4.18)$$

Note the symmetry breaking comes in as a result of  $\tilde{W}_{s,k_y,k_x=k_{x0}}^L$ , with the denominator acting as a quenching effect. Here  $\tilde{W}_{s,k_y,k_x}^L$  is the quasilinear weight defined by TGLF, which is evaluated at a single  $k_x$  rather than a sum over  $k_x$ . The two constants are  $\alpha_x = 1.21$  and  $\sigma_x = 2$ ,  $k_{x0} = 0.32k_y (k_{\text{max}}/k_y)^{0.7} (\gamma_E/\gamma_{\text{max}})$  and  $k_x^{\text{model}}$  is given by equation 4.15.

#### 4.4.2 QuaLiKiz

QuaLiKiz does not solve a gyrofluid system, but instead a simplified linear gyrokinetic system, born from the linear stability code Kinezero [125, 126]. It makes a number of assumptions regarding the solution to the linear stability to bring the calculation time down to those required for integrated modelling. These include the assumption that the linear potential

eigenfunctions take the form of a shifted Gaussian, for which the shift and width are solved. Other key differences to TGLF include only using large aspect-ratio shifted circle geometry<sup>6</sup>, used to simplify the calculation of linear integrals, a simplified Krook collision operator for trapped electrons, and the assumption of no magnetic fluctuations, negating the need for the calculation of the fluctuating Ampère’s law. A comprehensive overview of all aspects of the QuaLiKiz model is provided in [128].

### QuaLiKiz saturation rule

QuaLiKiz contains only one saturation rule [128, 129] which will now be outlined. It is fundamentally based on the mixing length argument and has a relatively simple spectral shape, covering both ion and electron scales. Initially assuming only the dominant mode is calculated, the saturation rule is constructed by first calculating the spectrum of  $\gamma_{k_y}/k_{\perp,Q}^2$ , where  $k_{\perp,Q}^2$  is an analytic approximation of the form of the squared perpendicular averaged over the eigenfunction, including a contribution that has been tuned to nonlinear gyrokinetic results. The point at which this ratio is maximised is defined as  $k_{\max,Q}$ , which is used to inform the spectral shape<sup>7</sup>.

$$\left. \frac{d}{dk_y} \left( \frac{\gamma_{k_y}}{k_{\perp,Q}^2} \right) \right|_{k_y=k_{\max,Q}} = 0 \quad (4.19)$$

The saturation rule has the form<sup>8</sup>

$$\left| \hat{\phi}_{k_y} \right|_{\text{QLK}}^2 = \left| \hat{\phi}_{k_y=k_{\max,Q}} \right|^2 S_{k_y} \quad (4.20)$$

where  $S_{k_y}$  is a normalised function describing the spectral shape, given by

$$S_{k_y} = \begin{cases} k_y/k_{\max,Q} & k_y \leq k_{\max,Q} \\ (k_y/k_{\max,Q})^{-3} & k_y > k_{\max,Q} \end{cases} \quad (4.21)$$

and  $\left| \hat{\phi}_{k_y=k_{\max,Q}} \right|^2$  is the value of the potential at the peak of the spectrum. Making use of the mixing length rule, this is modelled as

$$\left| \hat{\phi}_{k_y=k_{\max,Q}} \right|^2 = C_{\text{NL}} \left. \frac{\gamma_{k_y}}{k_{\perp,Q}^2} \right|_{k_{\max,Q}} \frac{T_e B}{k_{\max,Q} R_0 e} \quad (4.22)$$

---

<sup>6</sup>This is  $s - \alpha$  geometry [127].

<sup>7</sup>Note, this is generally not the same as  $k_{\max}$  defined for TGLF.

<sup>8</sup>Some works denote the binormal wavenumber  $k_{\theta}$  for historical reasons, even in the case that the parallel coordinate of the field-aligned system is labelled  $\theta$ . This wavenumber is defined  $k_{\theta} = qk_{\alpha}/r$ , and so coincides with our definition of  $k_y$ , equation 3.187.

where  $C_{\text{NL}}$  is a dimensionless prefactor that has been fitted to nonlinear gyrokinetic simulations. Its value depends on whether the binormal wavenumber under consideration is in the ITG scale or the ETG scale, a boundary taken at  $k_y \rho_i = 2$ ,

$$C_{\text{NL}} = \begin{cases} 271/s_{\text{fac}} & k_y \rho_i < 2 \text{ (ITG)} \\ 122 f_{\text{multi-scale}}/s_{\text{fac}} & k_y \rho_i > 2 \text{ (ETG)} \end{cases} \quad (4.23)$$

where  $s_{\text{fac}}$  is a factor introduced to better model cases of low magnetic shear [130]

$$s_{\text{fac}} = \begin{cases} 2.5 (1 - |\hat{s}|) & |\hat{s}| < 0.6 \\ 1 & |\hat{s}| > 0.6 \end{cases} \quad (4.24)$$

and  $f_{\text{multi-scale}}$  is a function which models the transition from ion-scale modes to electron-scale modes [53]

$$f_{\text{multi-scale}} = \left[ 1 + e^{-\frac{1}{5} \left( \frac{\gamma_{\text{ETG,max}}}{\gamma_{\text{ITG,max}}} - \sqrt{\frac{m_i}{m_e}} \right)} \right]^{-1} \quad (4.25)$$

where the growth rates correspond to the maxima of their respective spectra. Note these are examples of quasilinear model development, changes made after the conception of the model to better match nonlinear gyrokinetic results.

In the case of subdominant modes, the model for the saturated potential of mode  $j$  generalises to [129]

$$\left| \hat{\phi}_{k_y}^j \right|_{\text{QLK},j}^2 = C_{\text{NL}} \frac{\gamma_{k_y,j}}{k_{\perp,Q}^2} \Big|_{k_{\text{max},Q,j}} \frac{T_e B}{k_{\text{max},Q,j} R_0 e} S_{k_y} \frac{\gamma_{k_y,j}}{\gamma_{\text{max},k_y}} \quad (4.26)$$

where  $\gamma_{k_y,j}$  is the growth rate of the mode  $j$  at a given  $k_y$ ,  $k_{\text{max},Q,j}$  is the wavenumber at which the analogue of equation 4.19 for  $\gamma_{k_y,j}$  is satisfied, and  $\gamma_{\text{max},k_y}$  is the largest growth rate across all  $j$  for a given  $k_y$ . Note that in the case of a single eigenmode, the form of equation 4.26 reduces to that given by equations 4.20 and 4.22, as expected.

Now that the gyrokinetic and quasilinear theories of turbulence have been established we turn to the main purpose of this work, namely the performance of quasilinear models in isotopes differing from deuterium. In the interest of generality, reduced models will be referred to using the taxonomy  $\langle \text{linear inputs} \rangle - \langle \text{saturation rule} \rangle$ , to clearly specify the constituents of each model. For example TGLF-SAT2 refers to the fast model type of SAT2 with the reduced linear inputs of TGLF and CGYRO-SAT1 refers to the exact model type of linear CGYRO being input to SAT1.

## 4.5 Isotope scaling in quasilinear models

### 4.5.1 GyroBohm scaling in local gyrokinetic systems

We first examine how the gyroBohm arguments suggest various quantities of interest should scale with isotope mass from their ordering assumptions. Given that the ion gyroradius scales as  $\rho_i \propto \sqrt{A}$  and the thermal velocity scales as  $v_{\text{th},i} \propto 1/\sqrt{A}$ , we have

- Growth rates and frequencies  $\sim \mathcal{O}(v_{\text{th},i}/a) \implies \gamma_{\text{GB}}, \tilde{\omega}_{\text{GB}} \propto 1/\sqrt{A}$
- Fluxes (from section 2.6)  $\implies \Gamma_{\text{GB}}, Q_{\text{GB}} \propto \sqrt{A}$
- Saturated potentials  $\sim \mathcal{O}(\rho_* T/e) \implies \phi_{\text{GB}} \propto \sqrt{A}$
- Weight spectra  $W_{\text{GB}} \sim Q_{\text{GB}}/\phi_{\text{GB}}^2 \implies W_{\text{GB}} \propto 1/\sqrt{A}$

These collectively are the results of gyroBohm scaling, and are found to hold in sufficiently simple simulations, primarily characterised by adiabatic electrons. The inclusion of kinetic electrons and other physical mechanisms cause scalings to deviate from these results, and thus these scalings will be used as a reference to measure the deviations obtained from simulations.

### 4.5.2 Incorrect isotope scaling in quasilinear models

An example of the isotope effect as it appears in local nonlinear gyrokinetics was explicitly demonstrated in a paper by E. A. Belli *et al* [64], focusing on the role of kinetic electrons. It was shown using the gyrokinetic code CGYRO [38] that by increasing the equilibrium density gradient from a GA-standard case [99] baseline, one moves from a regime of ITG-dominated turbulence to TEM-dominated turbulence, accompanied by a reversal in the isotope scaling of the fluxes. It was suggested that this anti-gyroBohm scaling may not be captured by reduced turbulence models, in part due to the observation that the mixing length rule did not exhibit this isotope reversal.

To test the adherence between the nonlinear (NL) CGYRO results and the quasilinear model results, the NL CGYRO data from the density gradient scan presented in [64] is shown in figure 4.2, along with the data from equivalent simulations using TGLF-SAT1, TGLF-SAT2 and QuaLiKiz. Moving from  $a/L_n = 0.0$  to  $a/L_n = 3.0$ , one observes that the anti-gyroBohm scaling seen in the data from NL CGYRO for  $a/L_n > 2.0$  is not replicated in the results of any of the three quasilinear models. These all instead exhibit positive isotope scaling across the scan, indicating that the relevant physics to capture this isotope scaling reversal is missing.

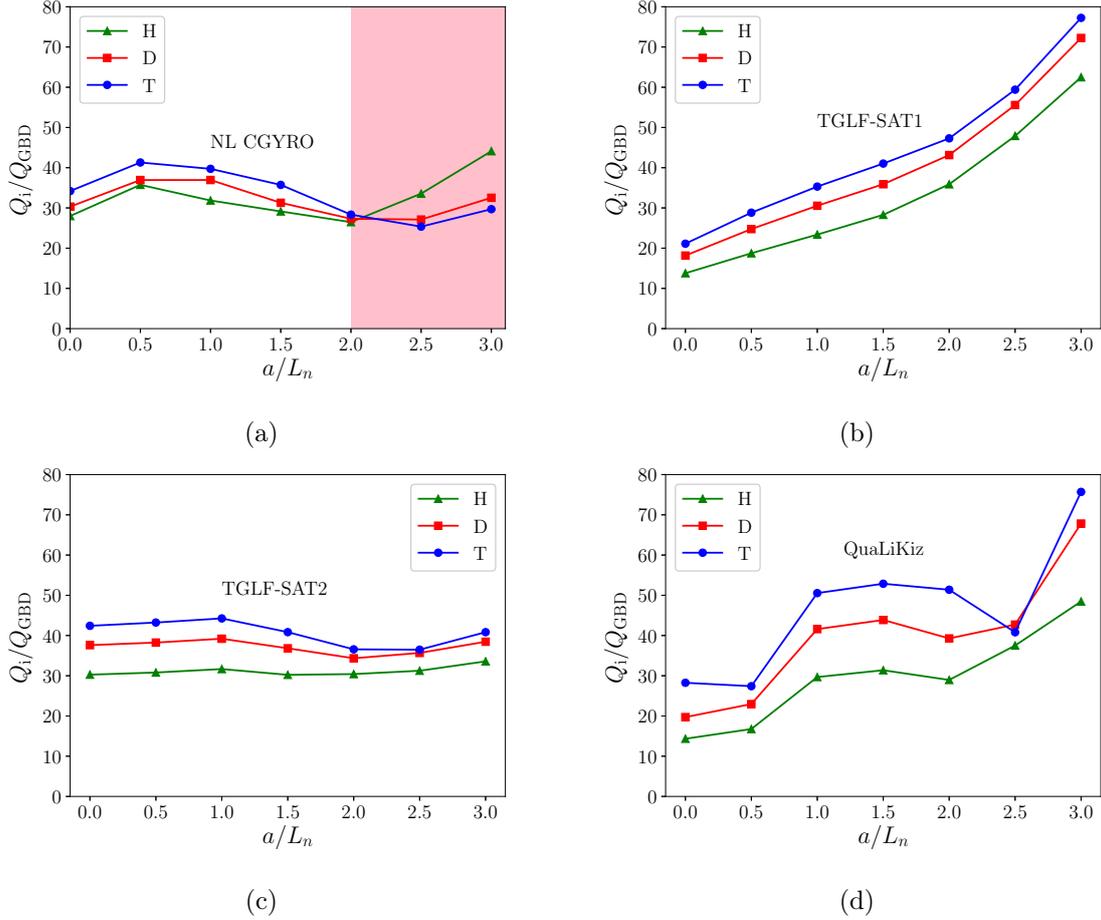


Figure 4.2: Turbulent ion energy fluxes against density gradient scale length for (a) NL CGYRO, (b) TGLF-SAT1, (c) TGLF-SAT2 and (d) QuaLiKiz in H, D and T, from a GA-std case baseline. Note that the isotope scaling reversal present in NL CGYRO (shaded region) is not recreated in any of the quasilinear models. Subfigure (a) reproduced from [64], with the permission of AIP Publishing.

To investigate and attempt to rectify the isotope scaling discrepancy observed in figure 4.2, we engage in a cycle of quasilinear model development, first generating a database of nonlinear gyrokinetic simulations against which to compare our quasilinear results.

## 4.6 Simulation database

A database of 43 nonlinear gyrokinetic simulations was generated using CGYRO. The database is primarily centered around the GA-std case, defined by  $a/L_{T_i} = a/L_{T_e} = 3.0$ ,  $a/L_n = 1.0$ ,  $T_i/T_e = 1.0$ ,  $\hat{s} = 1.0$ ,  $q = 2.0$ ,  $(a/c_s)\nu_{ee} = 0.1$  and circular Miller flux-surface geometry with  $r_M/R_{0M} = 1/6$ . Kinetic electrons and a single ion species are used for all cases. The three

isotopes that have been simulated are H, D and T, with  $m_i/m_D$  values of 0.5, 1.0 and 1.5 respectively. No rotation is included. All simulations are predominantly electrostatic, however include  $\delta A_{\parallel}$  fluctuations with a small plasma beta of  $\beta_{e,\text{unit}} = 0.05\%$  to allow for an increased time-step with negligible effect on the fluxes [38]. Additional parameters considered include the elongation  $\kappa_M$  and the Shafranov shift [131]  $\Delta = dR_{0M}/dr$ . Changes in temperature gradient scale lengths were kept constant between the ions and electrons ( $a/L_{T_i} = a/L_{T_e}$ ). The box size integer  $w$  has a value of 4 for all simulations except those changing the magnetic shear from its baseline value of  $\hat{s} = 1$ . For these cases, the box length integer was also changed so as to keep  $\hat{s}/w$  constant and thus the radial domain length unchanged (see equation 3.188). In table 4.1, the tokamak parameters that differ from the GA-std baseline are shown for the database.

Varied parameter	Values (label)	Fixed	Isotopes simulated
- (GA-std)	- (a)	-	H, D, T
$a/L_{T_i} = a/L_{T_e}$	1.5 (b), 2.25 (c), 3.5 (d)	-	H, D, T
$a/L_n$	2.0 (e), 3.0 (f)	-	H, D, T
$\hat{s}$	0.25 (g), 0.5 (h), 1.5 (i)	$\hat{s}/w = 1/4$	D
$(a/c_s)\nu_{ee}$	0.01 (j)	-	H, D, T
$(a/c_s)\nu_{ee}$	1.0 (k)	-	D
$T_i/T_e$	0.5 (l), 1.5 (m)	-	D
$q$	1.5 (n), 2.5 (o)	-	D
$\kappa_M$	1.25 (p), 1.5 (q), 2.0 (r)	-	D
$\Delta$	-0.125 (s), -0.25 (t), -0.5 (u)	-	D
$r_M/R_{0M}$	1/4 (v), 1/12 (w)	-	D
$(a/c_s)\nu_{ee}$	0.01 (x)	$a/L_n = 3.0$	H, D, T
$(a/c_s)\nu_{ee}$	0.05 (y)	$a/L_n = 3.0$	H, T
$(a/c_s)\nu_{ee}$	1.0 (z)	$a/L_n = 3.0$	D

Table 4.1: Details of the 43 nonlinear CGYRO cases that form the database. This set was generated to recreate a subset of the points of the density scan seen in figure 4.2, and then expanded out in a number of parameters of interest from the well-studied GA-std case to test the quasilinear models' dependence in these different areas. Note that the labels in the second column correspond with those in figure 5.8c, used for discerning the cases displayed in figures 5.8a, 5.8b, 5.12 and 5.13.

The resolutions used in this work for the nonlinear simulations are  $N_y = 40$  binormal modes,  $N_x = 224$  radial modes,  $N_\theta = 32$  parallel grid-points,  $N_\xi = 16$  pitch-angle grid-points and  $N_u = 8$  energy grid-points. The density of binormal modes is greater than that typically used in studies of similar cases [64], as it was found during convergence tests that this lower resolution can cause fluxes to be under-predicted (appendix E.3, see also [132]). All simulations were conducted at the ion scale up to  $k_y \rho_i = 1.0$ , where  $\rho_i = \sqrt{m_i/m_D} \rho_{\text{unit}}$ , to keep the radial and binormal domains constant relative to the main ion gyroradius. This binormal grid set-up was also used for the quasilinear model simulations of TGLF and QuaLiKiz shown in figure 4.2.

For each nonlinear case 39 linear simulations were conducted, corresponding to the 39 non-zero binormal modes present in their respective nonlinear simulations. An altered radial domain was used with  $N_x = 128$  and  $w = 1$ , due to the difference in grid requirements for the numerical convergence of the quantities of interest between nonlinear and linear runs.

A single nonlinear simulation of the GA-std case in D was repeated with  $N_x = 896$  keeping  $w = 4$ , for use in figure 5.3. This extended domain case plays no other part in the work.

Six additional linear simulations were performed scanning over relative concentrations of D and T in mixed D-T plasmas for the GA-std and  $a/L_n = 3.0$  cases in section 5.5.3. The values simulated are  $n_D/n_e = \{0.25, 0.5, 0.75\}$ , with  $n_T/n_e = 1 - n_D/n_e$ .

## 4.7 Model comparison and isotope scaling diagnosis

In this section, the results from the CGYRO database are compared with those obtained from CGYRO-SAT1, TGLF-SAT1 and QuaLiKiz-qlk. This is done to observe from where in the nonlinear gyrokinetic data the non-trivial isotope scaling originates, as per the flux breakdown in equation 4.7, as well as to identify why the current quasilinear models are not reproducing the NL CGYRO results. As an example case this section will primarily focus on the three  $a/L_n = 3.0$  simulations, due to the anti-gyroBohm scaling present in the fluxes.

### 4.7.1 Isotope scaling metric $\alpha_A$

To analyse the isotope scaling of these models the metric  $\alpha_A$  is introduced, such that for flux data in H, D and T, one may fit the data with a function of the form

$$f(A; C_A, \alpha_A) = C_A A^{\alpha_A} \tag{4.27}$$

where the values of  $C_A$  and  $\alpha_A$  are found via best fit to the data points<sup>9</sup>. An advantage to this metric is that one number  $\alpha_A$  can describe the scaling for three isotopes, whereas previously used metrics [54, 133] have been based around taking the ratio of two fluxes, and thus at least two numbers have been needed for an approximately equivalent description. The values of  $\alpha_A$  are also intuitive: in the case of fluxes, if one measures an  $\alpha_A$  value of  $\alpha_A \approx 0.5$ , one knows that the case follows approximate gyroBohm scaling. If one measures  $\alpha_A \approx 0.0$ , then the fluxes do not vary with isotope, and for  $\alpha_A < 0.0$ , the case exhibits anti-gyroBohm scaling. Outside of these limiting cases, one can quantify to what degree the case diverges from gyroBohm scaling: for two hypothetical cases of  $\alpha_A = 0.4$  and  $\alpha_A = 0.1$ , then both exhibit positive scaling, however the second ‘deviates more’ from gyroBohm than the first, by a defined quantitative amount.

In the case of only two isotopes being considered, the coefficients can be found analytically. For two isotopes  $A_1$  and  $A_2$ , we have  $f(A_1) = C_A A_1^{\alpha_A}$ ,  $f(A_2) = C_A A_2^{\alpha_A}$ , and hence

$$\alpha_A = \frac{\log(f(A_1)/f(A_2))}{\log(A_1/A_2)} \quad (4.28)$$

Any previous analysis involving ratios of quantities as a metric for isotope scaling can therefore be recast in terms of  $\alpha_A$ .

## 4.7.2 Flux comparison

### Total fluxes

Figure 4.3 shows the total ion energy fluxes for the GA-std case and the  $a/L_n = 3.0$  case against  $A$  for the four models, as well as TGLF-SAT2<sup>10</sup>. Fitted to each data set is the result of the metric fit with the measured values of  $\alpha_A$  displayed in the legend. Positive isotope scaling is seen in the GA-std case (figure 4.3a) for all 5 models, as one would expect for ITG-dominated turbulence representative of the quasilinear models’ training datasets<sup>11</sup>. QuaLiKiz remains close to the gyroBohm result.

<sup>9</sup> $C_A$  should approximately correspond to the value of  $f$  for the H isotope,  $A = 1$ .

<sup>10</sup>The results of TGLF-SAT2 are included in figure 4.3 only. The behaviour of SAT2 with isotope mass is very similar to that of SAT1, as demonstrated by the similarities in scaling metric between the fluxes of TGLF-SAT1 and TGLF-SAT2 in figure 4.3. This is in part due to the lack of consideration of isotopes other than deuterium in their development, as well as the similarity of their saturated potentials’ functional form. Beyond figure 4.3 therefore only the results of SAT1 are shown, with any conclusions regarding the isotope scaling of SAT1 also applicable to SAT2.

<sup>11</sup>The difference in magnitude between NL CGYRO and the quasilinear models is due to the increase in the NL CGYRO fluxes compared to previous datasets [52, 118], from the finer grid-resolutions used in this work.

For the  $a/L_n = 3.0$  case, anti-gyroBohm scaling is observed in the NL CGYRO data, whereas the three fast quasilinear models all continue to exhibit positive isotope scaling,  $\alpha_A > 0$ . For CGYRO-SAT1 one finds  $\alpha \approx 0.0$ , implying that the use of the linear solver of CGYRO compared with the reduced linear solver of TGLF is having an influence on the isotope scaling. As an exact linear solver type model, the resulting difference in isotope scaling of  $\Delta\alpha_A = -0.45$  between CGYRO-SAT1 and NL CGYRO can only come from either the QLA function, which is assumed to be constant in SAT1, and/or the functional form of the saturation rule.

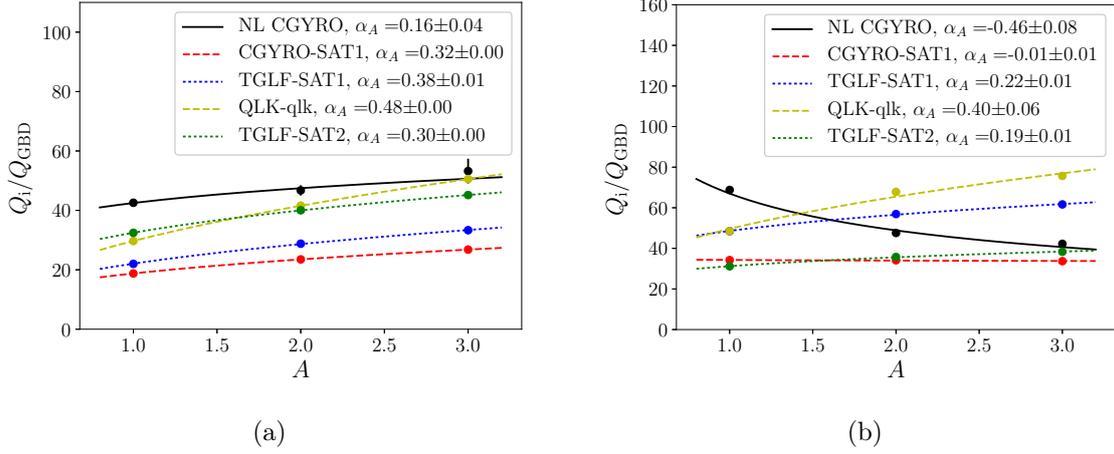


Figure 4.3: Turbulent ion energy flux against isotope mass for the GA-std case (a) and the  $a/L_n = 3.0$  case (b), for NL CGYRO (black), CGYRO-SAT1 (red), TGLF-SAT1 (blue), QuaLiKiz-qlk (yellow) and TGLF-SAT2 (green). Note the positive isotope scaling observed in both cases for the three fast quasilinear models, and the difference between the CGYRO-SAT1 and the TGLF-SAT1 results, originating solely from a difference in linear solver.

### Flux components

To probe this discrepancy further, the decomposition of the total fluxes into their flux components  $Q_{s,k_y}$  for the  $a/L_n = 3.0$  case is shown in figure 4.4a. Here the flux components for NL CGYRO are plotted against  $k_y\rho_i$  for the three isotope simulations, such that they have a shared  $k_y$ -axis. To quantify the isotope scaling of these flux components the isotope scaling metric  $\alpha_A$  can again be used, however now as a function of  $k_y\rho_i$ , to quantify how the scaling of the flux components varies across the spectrum. Hence for each value of  $k_y\rho_i$ , the three flux component data points are fitted using equation 4.27 with the resulting  $\alpha_A$  measurements forming an ‘ $\alpha_A$  line’, as shown in black in figure 4.4b. This exercise is repeated for CGYRO-SAT1, TGLF-SAT1 and QuaLiKiz-qlk. Also shown in figure 4.4b is a reference line at  $\alpha_A = 0.5$ , corresponding to the expected result if the flux components followed gyroBohm

scaling.

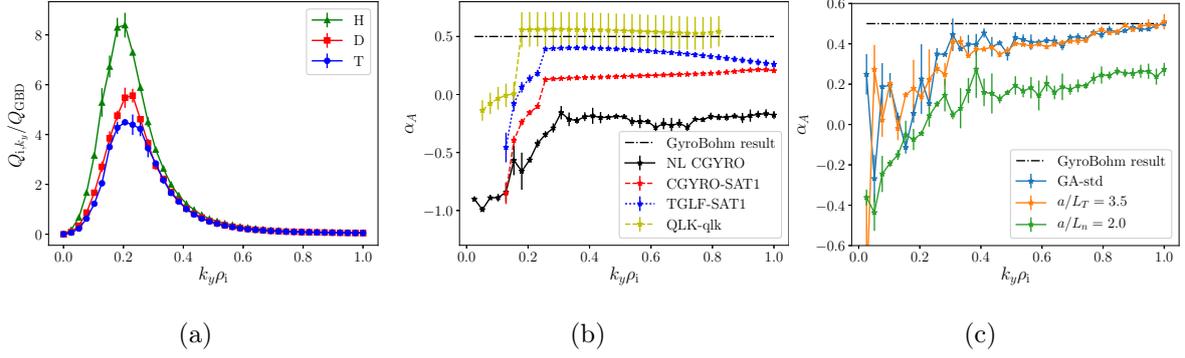


Figure 4.4: NL CGYRO ion flux components against wavenumber normalised to the main ion gyroradius  $k_y \rho_i$  for the  $a/L_n = 3.0$  case in H, D and T (a). In (b), the result of applying the  $\alpha_A$  metric at each  $k_y \rho_i$  is shown for the four models, as well as the value that would be seen if the fluxes followed gyroBohm scaling ( $\alpha_A = 0.5$ ). (c) Further isotope scaling results for NL CGYRO data from other cases in the database. The error bars shown for all models in (b) and (c) are the uncertainties in the fitted parameter  $\alpha_A$ .

For the NL CGYRO  $\alpha_A$  line, the isotope scaling is not uniform across the spectrum as one may expect, but instead two key features can be observed: an ‘offset’ from the gyroBohm value, most obviously seen in the region of  $k_y \rho_i > 0.3$ , and a variation in the isotope scaling with  $k_y \rho_i$  between  $0.0 < k_y \rho_i < 0.3$ . This observation is found to be general for the database, with a selection of other cases shown in figure 4.4c. It is also in this low  $k_y$  region that the flux component magnitudes are largest, and hence contribute most to the total flux and its isotope scaling. Because the larger  $k_y \rho_i$  region’s flux contribution is essentially negligible in comparison, figure 4.4 implies that to accurately capture the isotope scaling of the total flux, one must capture the shape of the flux component spectrum around the peak, as well as the isotope scaling characteristics observed in the low  $k_y$  region of the NL CGYRO data of figures 4.4b and 4.4c.

### 4.7.3 Comparison of flux constituents

Considering now the constituents of the flux components relevant to reduced models, as per equation 4.7, the two quantities in need of prescription are the QLA functions and the potentials. The results of analogous  $\alpha_A$  fitting exercises are shown in figures 4.5 and 4.6 for the ion energy QLA function and the saturated potentials respectively, along with their reference lines expected from gyroBohm scaling arguments. Looking at the  $\alpha_A$  line of the

QLA function in figure 4.5b, a small variation of  $\alpha_A \sim -0.1$  is seen in the region of low  $k_y$ . This is seen to be a general effect in figure 4.5c. Reduced models are not shown on these plots as they all assume the QLA function to be exactly 1, and so would be simply aligned with the gyroBohm-predicted result of  $\alpha_A = 0.0$ .

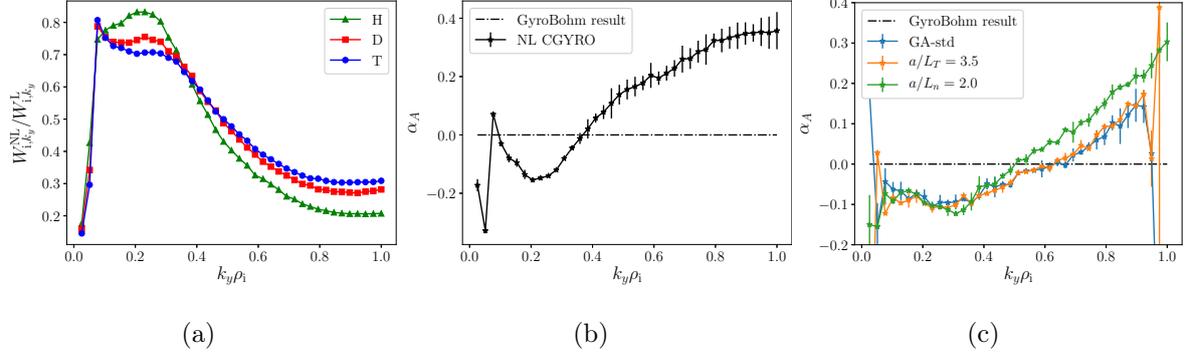


Figure 4.5: CGYRO ion energy QLA function against  $k_y \rho_i$  for the  $a/L_n = 3.0$  case in H, D and T (a). In (b), the result of applying the  $\alpha_A$  metric at each  $k_y$  is shown, including the reference value one would expect if gyroBohm scaling were followed ( $\alpha_A = 0.0$ ). This value is expected as the scaling of the linear and nonlinear weights is the same,  $W_{GB}^L, W_{GB}^{NL} \propto e^2 n_e c_s / T \propto A^{-0.5}$ , and thus their ratio cancels any predicted scaling dependence. The results of the quasilinear models are not shown as their QLA functions are taken to be constant, and thus exhibit no scaling ( $\alpha_A = 0.0$ ). (c)  $\alpha_A$  results of the QLA functions for additional cases in the database. The error bars shown for (b) and (c) are the uncertainties in the fitted parameter  $\alpha_A$ .

While there is some non-trivial isotope scaling in the low  $k_y$  region of the QLA function, this can be seen to be relatively small when compared to the difference between the NL CGYRO result and the gyroBohm-predicted scaling of the saturated potentials, shown in figures 4.6b and 4.6c. The crucial variation in isotope scaling in the low  $k_y$  region observed in the flux components is also seen to originate here. Hence the deviation from gyroBohm scaling of total fluxes in nonlinear gyrokinetics originates primarily in the saturated potentials, specifically in the region of low  $k_y$ .

Turning to the results of the current quasilinear models in figure 4.6b, TGLF-SAT1 and CGYRO-SAT1 both appear to recreate a portion of the isotope scaling variation in the low  $k_y$  region, indicating a reasonably accurate spectral shape for SAT1. This implies that it is their offsets that are primarily responsible for the total flux scalings seen in figure 4.3b, with the different values attributed to the difference in linear solver, and thus need improving. QuaLiKiz-qlk on the other hand does not exhibit this continuous variation, due to the

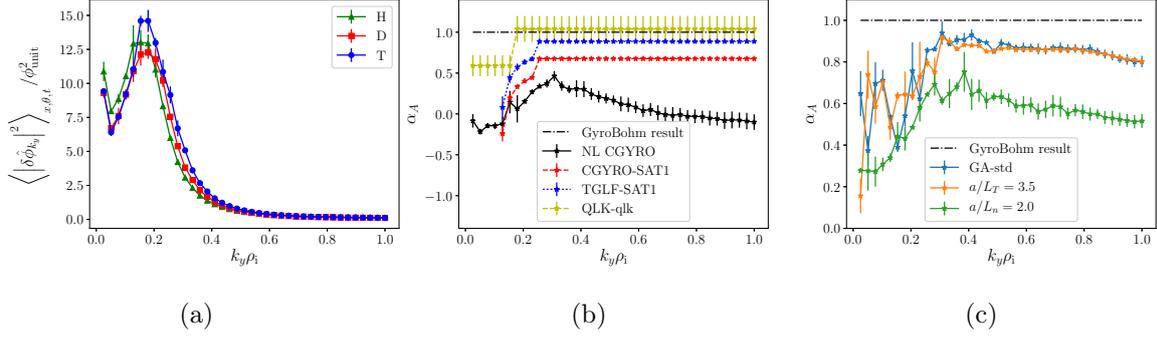


Figure 4.6: NL CGYRO squared potential magnitudes against  $k_y \rho_i$  for the  $a/L_n = 3.0$  case in H, D and T (a). In (b), the result of applying the  $\alpha_A$  metric at each  $k_y$  is shown, including the reference value one would expect if gyroBohm scaling were followed ( $\alpha_A = 1.0$ ), originating from  $\delta\phi_{\text{GB}}^2 \propto \rho_*^2 (T/e)^2 \propto A$ . (c) Isotope scaling of the saturated potentials from NL CGYRO for other cases in the database. The error bars shown for all models in (b) and (c) are the uncertainties in the fitted parameter  $\alpha_A$ .

comparatively simple functional form of its spectral shape. The majority of the scaling in the higher  $k_y$  region can also be seen to approximately lie at the gyroBohm level for the potentials. These results indicate that neither SAT1 or the QuaLiKiz saturation rule are fully capturing the relevant physics that describes the variation in the saturated potentials with isotope, whether as a consequence of missing the variation in the low  $k_y$  region, the correct offset from the gyroBohm result, or a combination of the two.

To summarise the findings of this section, then in order for quasilinear models to capture the isotope scaling of turbulent fluxes seen in nonlinear gyrokinetics, one must have a saturation rule that accurately predicts the spectral shape of the potentials around the peak, as well as captures the  $\alpha_A$  line characteristics observed in figures 4.6b and 4.6c, in particular the low  $k_y$  region as this dominates the flux in ion-scale turbulence. The model must therefore have sufficient functional complexity to capture the variation of the scaling with  $k_y$  in the low  $k_y$  region, and an accurate prediction of the offset. A new saturation rule, derived in light of these observations, will now be considered.

## Chapter 5

# Development of a new saturation rule

This chapter details the construction of the new saturation rule SAT3, which will for the first time recreate the properties of the isotope scaling seen in the previous chapter. A summary of the model is presented in Section 5.7.

The methodology applied in this work to quasilinear model development is a ‘bottom-up approach’, in which nonlinear gyrokinetic data is first analysed before prescribing relations between the linear physics and the saturated state, as opposed to prescribing an a priori theoretical rule. Each aspect of the quasilinear model will be considered in turn, namely the shape of the potential spectrum, the saturation level, as well as the validity of the quasilinear approximation.

In the calculation of quasilinear fluxes, saturation rules are only strictly required to predict the 1D grid-independent saturated potentials,  $\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t} / \Delta k_y$ , however in the interest of generality the two dimensional spectrum in  $k_y$  and  $k_x$  will first be considered,  $\langle |\delta\hat{\phi}_{k_x,k_y}|^2 \rangle_{\theta,t}$ . One can obtain the 1D potentials needed from the 2D spectrum simply by summing the potentials over  $k_x$  via

$$\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t} = \sum_{k_x} \langle |\delta\hat{\phi}_{k_x,k_y}|^2 \rangle_{\theta,t}. \quad (5.1)$$

### 5.1 2D spectrum

For all cases in the database and at all values of  $k_y$ , some common features were observed regarding the 2D potential slices in  $k_x$ , a representative example of which is shown in figure 5.1. All spectra considered are approximately even functions about a single peaked value, and tend to 0 for large  $|k_x|$ . Due to the absence of any symmetry-breaking effects [134] in the database, the spectra are always observed to peak at  $k_x = 0$ . It has been shown however that

in at least the case of non-zero  $E \times B$  shear a shift in the peak can be produced [112, 123, 135], and so a non-zero peak position  $K(k_y)$  is considered in the following.

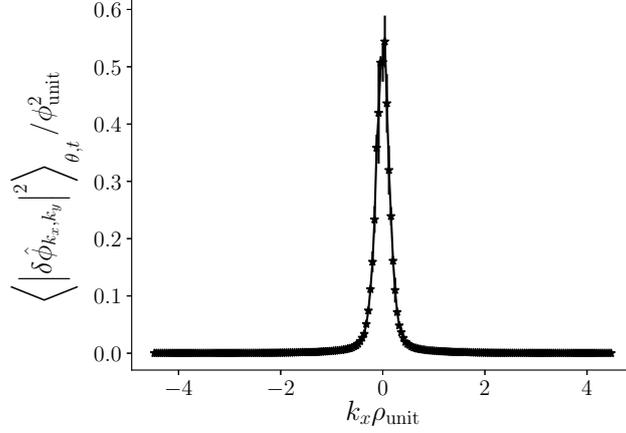


Figure 5.1: Example of a  $k_x$ -slice of the 2D potential magnitude spectrum, exhibiting the common features of an even, singularly peaked function with vanishing limits. Taken from the GA-std case in D, for  $k_y \rho_{\text{unit}} = 10/39 = 0.256$ .

The foregoing observations can all be accommodated into a functional description of the spectra by Taylor expanding the inverse of the potentials with the form

$$\frac{1}{\langle |\delta \hat{\phi}_{k_x, k_y}|^2 \rangle_{\theta, t}} = C_0(k_y) + (k_x - K(k_y))^2 C_1(k_y) + (k_x - K(k_y))^4 C_2(k_y) + \dots \quad (5.2)$$

where  $C_i(k_y)$  are  $k_y$ -dependent coefficients yet to be determined. Evaluating equation 5.2 at  $k_x = K$ , one finds  $C_0(k_y) = 1 / \langle |\delta \hat{\phi}_{k_x=K, k_y}|^2 \rangle_{\theta, t}$ , the inverse of the potentials at the peak. By truncating the expansion at  $\mathcal{O}(k_x^4)$ , taking out a factor of  $C_0$ , relabelling the coefficients  $C_i/C_0 \rightarrow C_i$  and inverting the equation one obtains a simple approximation for the shape of the potential spectrum,

$$\langle |\delta \hat{\phi}_{k_x, k_y}|^2 \rangle_{\theta, t} = \frac{\langle |\delta \hat{\phi}_{k_x=K, k_y}|^2 \rangle_{\theta, t}}{1 + C_1(k_y) (k_x - K(k_y))^2 + C_2(k_y) (k_x - K(k_y))^4}. \quad (5.3)$$

The physical interpretation of the coefficients  $C_1, C_2$  and  $K$  can be determined from considerations of the first three  $k_x$  moments of the 2D spectrum. The zeroth order moment is simply the 1D potential, given by equation 5.1. The first and second order moments define the mean  $k_x$  value  $\langle k_x \rangle$  and the radial width of the spectrum  $\left( \sigma_{k_y} = \sqrt{\langle k_x^2 \rangle - \langle k_x \rangle^2} \right)$  respectively,

$$\langle k_x \rangle = \frac{\sum_{k_x} k_x \langle |\delta \hat{\phi}_{k_x, k_y}|^2 \rangle_{\theta, t}}{\langle |\delta \hat{\phi}_{k_y}|^2 \rangle_{x, \theta, t}} \quad (5.4)$$

$$\sigma_{k_y}^2 = \frac{\sum_{k_x} (k_x - \langle k_x \rangle)^2 \left\langle \left| \delta \hat{\phi}_{k_x, k_y} \right|^2 \right\rangle_{\theta, t}}{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x, \theta, t}}. \quad (5.5)$$

By approximating the summations over  $k_x$  in equations 5.1, 5.4 and 5.5 as integrals via  $\sum_{x=a}^b f(x_i) \Delta x \approx \int_a^b f(x) dx$  and using the expression given by equation 5.3 for the 2D potential spectrum, one can evaluate the resulting integrals over  $k_x$  analytically (appendix E.4). Assuming a sufficiently small  $\Delta k_x$  and a sufficiently large  $k_x$  range<sup>1</sup>, one can show

$$K = \langle k_x \rangle, \quad C_2(k_y) = \frac{1}{\sigma_{k_y}^4}, \quad C_1(k_y) = \left[ \left( \frac{\pi \left\langle \left| \delta \hat{\phi}_{k_x = \langle k_x \rangle, k_y} \right|^2 \right\rangle_{\theta, t} \sigma_{k_y}}{\Delta k_x \left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x, \theta, t}} \right)^2 - 2 \right] \frac{1}{\sigma_{k_y}^2} \quad (5.6)$$

and thus equation 5.3 becomes

$$\left\langle \left| \delta \hat{\phi}_{k_x, k_y} \right|^2 \right\rangle_{\theta, t} = \frac{\left\langle \left| \delta \hat{\phi}_{k_x = \langle k_x \rangle, k_y} \right|^2 \right\rangle_{\theta, t}}{1 + \left[ \left( \frac{\pi \left\langle \left| \delta \hat{\phi}_{k_x = \langle k_x \rangle, k_y} \right|^2 \right\rangle_{\theta, t} \sigma_{k_y}}{\Delta k_x \left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x, \theta, t}} \right)^2 - 2 \right] \left( \frac{k_x - \langle k_x \rangle}{\sigma_{k_y}} \right)^2 + \left( \frac{k_x - \langle k_x \rangle}{\sigma_{k_y}} \right)^4}. \quad (5.7)$$

The prefactor of the  $k_x^2$  term can be interpreted in relation to the normalised fall-off from the spectrum peak. Evaluating equation 5.7 at  $k_x = \langle k_x \rangle \pm \sigma_{k_y}$ , one finds:

$$\left( \frac{\pi \left\langle \left| \delta \hat{\phi}_{k_x = \langle k_x \rangle, k_y} \right|^2 \right\rangle_{\theta, t} \sigma_{k_y}}{\Delta k_x \left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x, \theta, t}} \right)^2 = \frac{\left\langle \left| \delta \hat{\phi}_{k_x = \langle k_x \rangle, k_y} \right|^2 \right\rangle_{\theta, t}}{\left\langle \left| \delta \hat{\phi}_{k_x = \langle k_x \rangle \pm \sigma_{k_y}, k_y} \right|^2 \right\rangle_{\theta, t}}. \quad (5.8)$$

Equation 5.7 is similar to that used in TGLF's SAT1 and SAT2 [52, 118], in which the potentials are modelled as a squared Lorentzian, which assumes  $\sigma_{k_y}^2 C_1(k_y) = 2$  for all  $k_y$  in all cases<sup>2</sup>. Equation 5.7 generalises this assumption, retaining a degree of freedom in the description of the fall-off of the spectrum. The measured NL CGYRO values of  $\sqrt{\sigma_{k_y}^2 C_1(k_y)}$  for the GA-standard case in D are shown in figure 5.2 exhibiting strong variation against  $k_y$ , particularly in the low  $k_y$  region, indicating the value of this generalisation.

The quality of adherence to the data for equation 5.7 is demonstrated in figures 5.3a and 5.3b, using the radially-extended GA-std D simulation<sup>3</sup>. Calculating the moments from the raw

<sup>1</sup>This can be shown to require  $k_{x, \text{lim}} \gg |\langle k_x \rangle| + \sigma_{k_y}$ , for range  $\{-k_{x, \text{lim}}, k_{x, \text{lim}}\}$ .

<sup>2</sup>The values of  $\sigma_{k_y}$  for SAT1 and SAT2 are obtained via best-fit close to the peak of the data rather than calculation, and so can be considered an 'effective width' if the spectrum *were* a squared Lorentzian.

<sup>3</sup>This case has  $N_x = 896$  for the same  $(k_y \rho_{\text{unit}})_{\text{min}}$  and box-size integer  $w$ , and so contains four times as many  $k_x$  values for the same minimum  $k_x$ .

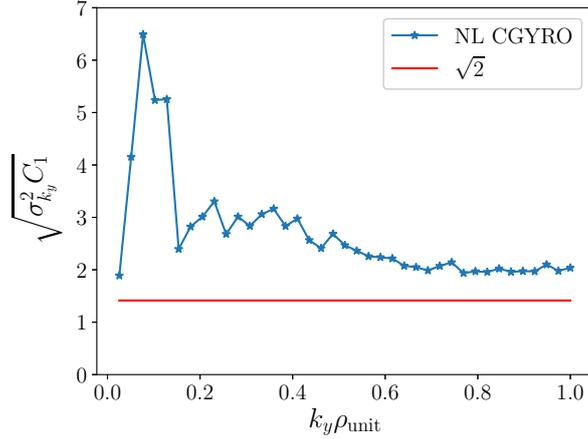


Figure 5.2: Measured values of  $\sqrt{\sigma_{k_y}^2 C_1}$  against  $k_y \rho_{\text{unit}}$  for the radially-extended ( $N_x = 896$ ) GA-standard case simulation in D, with  $\sqrt{\sigma_{k_y}^2 C_1} = \sqrt{2}$  marked in red, corresponding to the squared Lorentzian assumption.

data, it can be observed that the functional form holds extremely well over a wide domain and many orders of magnitude in range. To explore the different cascade regimes predicted by this equation, the simplifying notation here is introduced such that equation 5.7 is written:

$$P(X) = \frac{P_0}{1 + CX^2 + X^4} \quad (5.9)$$

where  $P = \left\langle \left| \delta \hat{\phi}_{k_x, k_y} \right|^2 \right\rangle_{\theta, t}$ ,  $X = |(k_x - \langle k_x \rangle) / \sigma_{k_y}|$ ,  $P_0 = \left\langle \left| \delta \hat{\phi}_{k_x = \langle k_x \rangle, k_y} \right|^2 \right\rangle_{\theta, t}$  and  $C = \sigma_{k_y}^2 C_1(k_y)$ . Assuming  $\sqrt{C} \gg 1$ , three distinct regions of scaling are predicted<sup>4</sup>:

$$P \approx \begin{cases} P_0 & X \ll \frac{1}{\sqrt{C}} \\ \frac{P_0}{CX^2} & \frac{1}{\sqrt{C}} \ll X \ll \sqrt{C} \\ \frac{P_0}{X^4} & X \gg \sqrt{C} \end{cases} \quad (5.10)$$

which are seen to be present in the data, as evidenced by the model-data agreement between the limiting regions marked in 5.3b.

Having modelled the  $k_x$ -dependence for the spectrum, the  $k_y$ -dependence will now be considered. Four as-yet unmodelled  $k_y$ -dependent quantities are present in equation 5.7: The 1D potential  $\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x, \theta, t}$ , the spectrum peak  $\left\langle \left| \delta \hat{\phi}_{k_x = \langle k_x \rangle, k_y} \right|^2 \right\rangle_{\theta, t}$ , the peak position  $\langle k_x \rangle$  and the radial spectral width  $\sigma_{k_y}$ . In order to have a full 2D potential model, one must uniquely constrain these quantities by providing four equations describing them. This is the chosen method of TGLF SAT1 [118], for which one of the four equations is the squared Lorentzian

<sup>4</sup>As one reduces the value of  $\sqrt{C}$  to  $\sim 1$ , and further to  $\sqrt{C} \ll 1$ , the  $P_0/CX^2$  scaling is found to be suppressed.

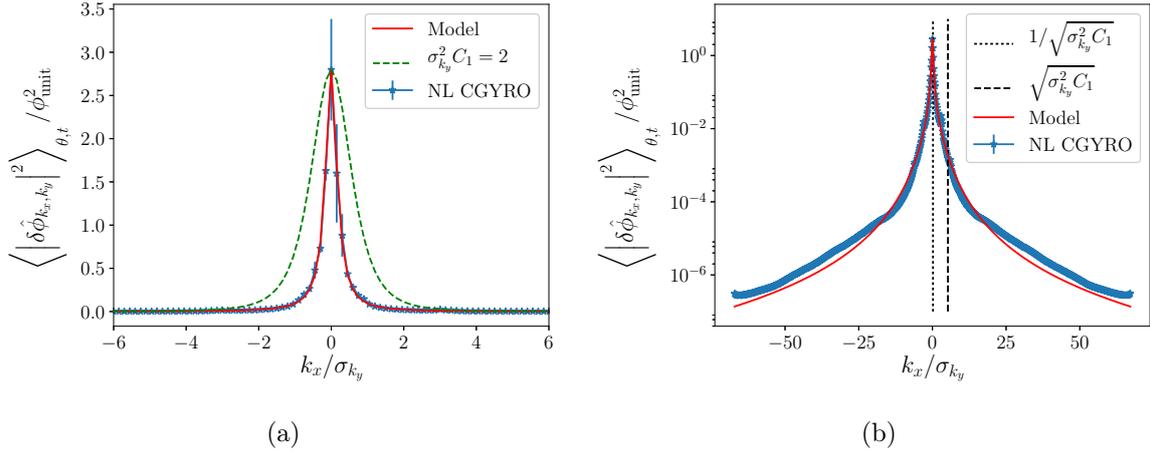


Figure 5.3: (a) An example of equation 5.7 (red) applied to NL CGYRO data (blue) for  $k_y \rho_{\text{unit}} = 5/39 = 0.128$  of the radially-extended GA-std case, plotted against  $k_x / \sigma_{k_y}$ . The squared Lorentzian model with the spectral width taken from the raw data is shown in green. (b) The same spectrum as (a) shown on a logarithmic  $y$ -axis, with the predicted scaling regime limits marked.

assumption. However, flux calculations ultimately only use the 1D potential,  $\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x, \theta, t}$ , and so technically this is the only quantity that needs to be prescribed for a reduced model. This potential could be modelled directly as a 1D function of  $k_y$ , such as is done by QuaLiKiz [130], but attempting to relate the remaining spectral quantities to one another first can help to reveal structure in the nonlinear saturated state, which can be exploited to build a saturation rule with a higher fidelity physical basis.

## 5.2 Saturation equations

### 5.2.1 Radial spectral width parameterisation

The following relation was discovered predominantly via empirical experimentation with the nonlinear database, and describes a seemingly fundamental relation between the zeroth radial moment and the second order radial moment of the 2D spectrum. It is observed that the zeroth moment is very well modelled by the equation

$$\frac{\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x, \theta, t}}{\Delta k_y} = c_0 \sigma_{k_y}^{c_1} \quad (5.11)$$

for  $k_y > 0$ , where  $c_0$  and  $c_1$  are case-dependent parameters. When allowing these parameters to be fitted to the NL CGYRO data, as shown for various cases in figure 5.4, it can be seen

that the two curves adhere to one another extremely closely<sup>5</sup>.

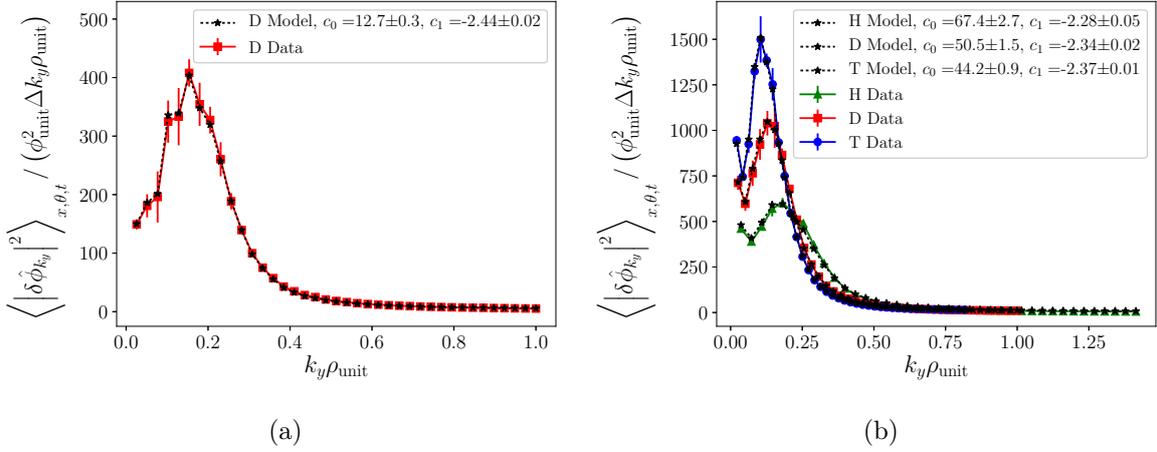


Figure 5.4: (a) The model of equation 5.11 (black) applied to the NL CGYRO grid-independent potentials of the GA-standard case in D (red), allowing the two parameters to be fitted to the data. (b) Further examples of equation 5.11, for the  $a/L_n = 3.0$ ,  $\nu_{ee}(a/c_s) = 0.01$  case in H, D and T.

For each case, this constitutes only two fitted parameters for 39  $k_y$  values, rendering there no question of over-fitting. Moreover, the measured value of the exponent  $c_1$  is found to be strongly consistent across cases, as displayed in figure 5.5, with an approximate value of  $c_1 = -2.42$  for the database, implying equation 5.11 resembles something close to a seemingly universal conservation law.

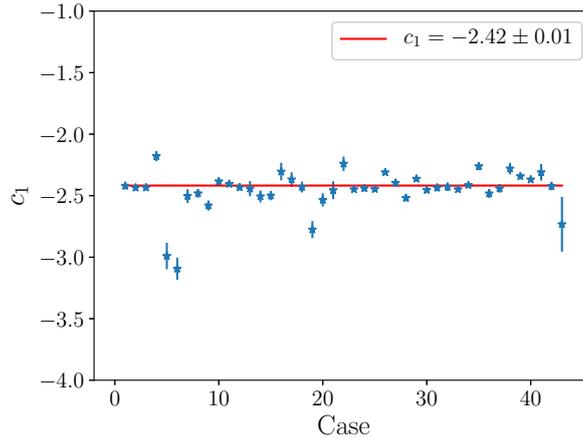


Figure 5.5: Measured values of the  $c_1$  parameter across all nonlinear simulations in the database. The error bars shown are the uncertainties in the fitted parameter  $c_1$ .

This observation is perhaps the most important point of this work, as it appears to describe

<sup>5</sup>For reference, example data for  $\sigma_{k_y} \rho_{\text{unit}}$  corresponding to the case of figure 5.4a is shown in figure 5.6.

a general and robust relation between two hypothetically-independent moments of the 2D potential spectrum. Somewhat counter-intuitively it suggests that the area under the radial spectrum, related to the zeroth moment, is independent of its peak value, and instead only depends on the width. This observed relation is understood to be novel, and a physical mechanism for why this is the case is yet to be put forward. It is this observation that forms the core of the new saturation rule.

### 5.2.2 Parameter considerations

Given sufficient ability to model the case-dependent values of  $c_0$ , equation 5.11 provides one of the four equations necessary to uniquely define the 2D potential spectrum. However, for every parameter introduced by such equations, a linear-physics based model for its case-dependent calculation will be required. Because equation 5.11 implies that the 1D potential is independent of the peak spectrum value and  $\langle k_x \rangle$ , the choice is made in this work to model  $\sigma_{k_y}$  only. Doing so will leave the 2D spectrum with an arbitrary peak value and location, but will have both its zeroth and second moments defined.

### 5.2.3 Model for the radial spectral width

To inform the model for  $\sigma_{k_y}$ , a representative plot is shown in figure 5.6. Equation 5.11 implies two important qualities of  $\sigma_{k_y}$ : firstly, that the characteristic peak observed in the 1D potentials should correspond to a minimum in  $\sigma_{k_y}$  at the same position in  $k_y$ , which is indeed seen, and is denoted  $k_{\min}$ . This is a novel observation of this work, in part due to the well-resolved fluctuation averages obtained from extended simulation times and the increased density of binormal grid-points used in the database. The second quality is that, because the values of the potentials away from the peak are comparatively small and contribute negligibly to the overall flux, the accurate modelling of  $\sigma_{k_y}$  in this higher  $k_y$  region is not required for successful flux prediction.

To incorporate the observed minimum into the model a quadratic polynomial is used. A quadratic over the entire  $k_y$  domain capturing this minimum can however become too large in the middle  $k_y$  region, and so a piecewise function is constructed, with a first-order polynomial used past a certain point,  $k_P$ . This is taken to be  $k_P = 2k_{\min}$ . The two regions are connected

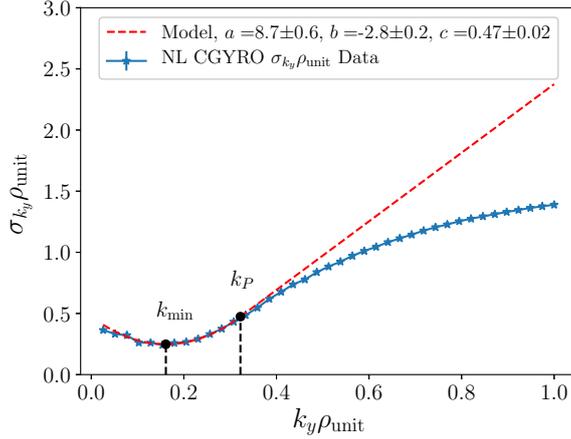


Figure 5.6: Example of  $\sigma_{k_y} \rho_{\text{unit}}$  data obtained for the GA-standard case in D (blue). In red is the model for the widths given by equation 5.12, with the parameters  $a, b$  and  $c$  fitted to the data. Also marked are  $k_{\text{min}}$  and  $k_P$ , the positions of the minimum and piecewise connection point respectively.

by imposing continuity of the function, as well as continuity of the gradient<sup>6</sup>, at  $k_y = k_P$ :

$$\sigma_{k_y} = \begin{cases} ak_y^2 + bk_y + c & 0 < k_y \leq k_P \\ (2ak_P + b)k_y + c - ak_P^2 & k_P < k_y < \infty \end{cases} \quad (5.12)$$

where  $a, b$  and  $c$  are coefficients to be modelled. Here the function is assumed to extend to infinity.

By combining equations 5.11 and 5.12, one obtains an equation for the 1D potentials solely as a function of  $k_y$  and the parameters  $\{c_0, c_1, a, b, c\}$ . To make progress in modelling these parameters, they shall first be defined in terms of more physically meaningful quantities.

### 5.3 Physical interpretation of saturation rule parameters

The exponent  $c_1$  is dimensionless and has already been fitted to the nonlinear database in figure 5.5. An expression for  $c_0$  can be obtained by evaluating equation 5.11 at some given point  $k_0$ , to be determined:

$$c_0 = \frac{\left\langle \left| \delta \hat{\phi}_{k_y=k_0} \right|^2 \right\rangle_{x,\theta,t}}{\Delta k_y} \left( \frac{1}{\sigma_{k_y=k_0}} \right)^{c_1} \quad (5.13)$$

<sup>6</sup>As noted, the accurate modelling of the higher  $k_y$  region is nonessential for flux prediction. Fixing the gradient causes the second branch of equation 5.12 to become a first-order Taylor expansion extending past  $k_P$ ,  $\sigma_{k_y}(k_P + \Delta k_y) = \sigma_{k_y}(k_P) + \Delta k_y \left( d\sigma_{k_y}/dk_y|_{k_y=k_P} \right)$ , which is sufficient to capture the non-negligible contribution to the flux.

giving the overall saturation model as

$$\frac{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}}{\Delta k_y} = \frac{\left\langle \left| \delta \hat{\phi}_{k_y=k_0} \right|^2 \right\rangle_{x,\theta,t}}{\Delta k_y} \left( \frac{\sigma_{k_y}}{\sigma_{k_y=k_0}} \right)^{c_1} \quad (5.14)$$

Equation 5.14 expresses the 1D potentials as a product of their magnitude at a given point  $k_0$ ,  $\left\langle \left| \delta \hat{\phi}_{k_y=k_0} \right|^2 \right\rangle_{x,\theta,t} / \Delta k_y$  and a  $k_y$ -dependent function describing the shape of the spectrum,  $(\sigma_{k_y} / \sigma_{k_y=k_0})^{c_1}$ . These quantities define the saturation level and the spectral shape respectively. The variation in isotope scaling observed with  $k_y$ , for example shown in figure 4.6b, comes entirely from the spectral shape, with the offset attributed to the saturation level.

As the radial spectral widths appear only in the above normalised form for the saturation model, a reduction in the degrees of freedom is provided. Assuming  $0 < k_0 \leq k_P$ , one can divide through by one of the expansion coefficients in the spectral widths, chosen here to be  $b$ :

$$\frac{\sigma_{k_y}}{\sigma_{k_y=k_0}} = \begin{cases} \left( \frac{a}{b} k_y^2 + k_y + \frac{c}{b} \right) / \left( \frac{a}{b} k_0^2 + k_0 + \frac{c}{b} \right) & 0 < k_y \leq k_P \\ \left( \left( 2 \frac{a}{b} k_P + 1 \right) k_y + \frac{c}{b} - \frac{a}{b} k_P^2 \right) / \left( \frac{a}{b} k_0^2 + k_0 + \frac{c}{b} \right) & k_P < k_y < \infty \end{cases} \quad (5.15)$$

and so the number of unknown coefficients reduces from 3 to 2, now requiring only the ratios  $a/b$  and  $c/b$ . A more transparent physical interpretation of  $a/b$  can be obtained by considering the definition of the minimum of the spectral widths. Imposing a minimum at  $k_{\min}$  in equation 5.12, one finds

$$\left. \frac{d\sigma_{k_y}}{dk_y} \right|_{k_y=k_{\min}} = 2ak_{\min} + b = 0 \quad (5.16)$$

and therefore  $a/b = -1/(2k_{\min})$ . With the coefficients now recast, this leaves 3 quantities to be modelled, for a given choice of  $k_0$ :  $k_{\min}$ ,  $c/b$ , and  $\left\langle \left| \delta \hat{\phi}_{k_y=k_0} \right|^2 \right\rangle_{x,\theta,t} / \Delta k_y$ .

Note, as of yet no appeals to linear physics have been made, the model derived thus far has come solely from considerations of nonlinear gyrokinetic spectra. By taking this bottom-up approach, the validity of the underlying functional forms of the saturation rule equations is guaranteed, up to the hypothetical quality observed when the parameters are fitted to the data. This foundation implies that as linear physics based approximations for the three quantities above improve, the model will tend towards being as accurate as when fitted in this way.

## 5.4 Approximating the parameters from linear physics quantities

The new saturation rule parameters will now be modelled entirely from linear physics. Following from the zonal flow mixing arguments used in SAT1 and SAT2, described in section 4.4.1, here we take  $k_{\max}$  and  $\gamma_{\max}$  to be a characteristic wavenumber and frequency of the turbulent saturation for SAT3. We therefore build linear models for  $k_{\min}$ ,  $c/b$ , and  $\left\langle \left| \delta \hat{\phi}_{k_y=k_0} \right|^2 \right\rangle_{x,\theta,t} / \Delta k_y$  by assuming proportionality to the dimensionally-consistent combination of  $\gamma_{\max}$  and  $k_{\max}$ , as well as the equilibrium quantity  $B_{\text{unit}}$ .

Considering first the dimensionality of  $b$  and  $c$  individually from equation 5.12, one finds  $b$  to be dimensionless and  $c$  to have dimensions of  $k_y$ , giving  $c/b$  dimensions of  $k_y$ . The quantity  $k_{\min}$  also has dimensions of  $k_y$ , and so per the method above these are both modelled as proportional to  $k_{\max}$ . The database-fitting exercise is carried out in figure 5.7, resulting in  $c/b = -0.751k_{\max}$  and  $k_{\min} = 0.685k_{\max}$ .

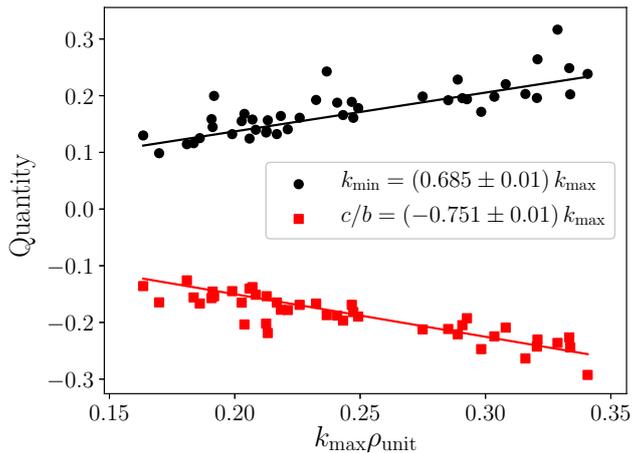


Figure 5.7: Fitting exercise of the two quantities modelled as proportional to  $k_{\max}$  across the NL CGYRO database:  $k_{\min}$  (black circles), and  $c/b$  (red squares), used in defining the model for the spectral shape  $\sigma_{k_y} / \sigma_{k_y=k_0}$  (equation 5.15).

For the saturation level  $\left\langle \left| \delta \hat{\phi}_{k_y=k_0} \right|^2 \right\rangle_{x,\theta,t} / \Delta k_y$ , the value of  $k_0$  is chosen to be  $k_0 = 0.6k_{\min}$ , as it is at this position that the model scatter for the following is found to be minimised. The dimensionally-consistent combination of linear quantities for the saturation level is  $\propto B_{\text{unit}}^2 \gamma_{\max}^2 / k_{\max}^5$ . Note that this is a similar form to that derived from considerations of balance between linear growth and turbulent  $E \times B$  advection [120].

When tested against the database,  $B_{\text{unit}}^2 \gamma_{\max}^2 / k_{\max}^5$  is found to be a good model for the cases in which the dominant mode is the ITG (figure 5.8a). However, for those with TEMs present,

$B_{\text{unit}}^2 \gamma_{\text{max}}^2 / k_{\text{max}}^4$  in dimensionless units is found to be in much better agreement, indicating a difference in physical saturation between the two mode types (figure 5.8b). The TEM saturation in dimensional units is therefore assumed proportional to  $\gamma_{\text{max}}^2 \rho_{\text{unit}} B_{\text{unit}}^2 / k_{\text{max}}^4$ . It is here explicitly that a large cause of the difference in isotope scaling between the two mode types can be seen, and the reason why previous saturation models have failed to recreate the isotope scalings seen in TEM-dominated turbulence.

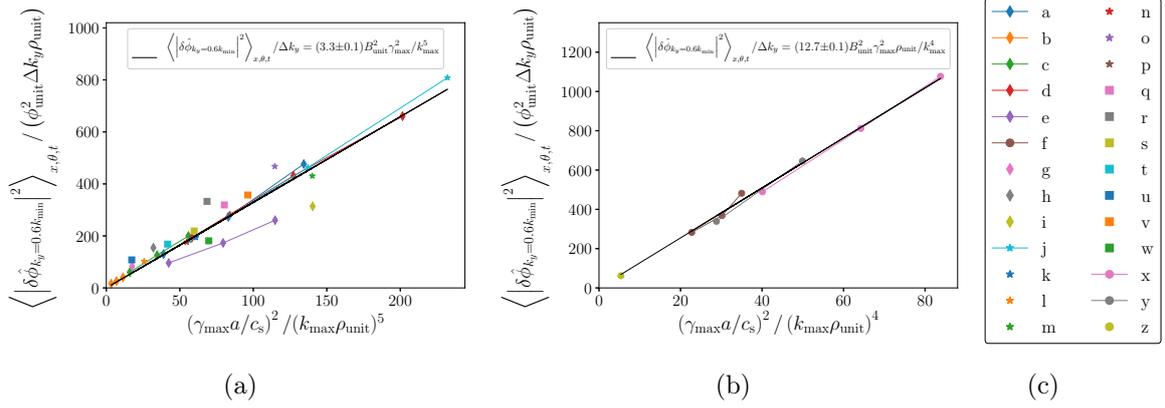


Figure 5.8: Potential magnitudes evaluated at  $0.6k_{\min}$  against a combination of linear quantities ((a) and (b)). Subfigure (c) displays the marker type used for each case, with the labels corresponding to those shown in table 4.1. Data points of the same colour connected by a line indicate simulations of the same equilibrium but of different isotopes. In (a), the cases with a dominant ITG instability scale approximately with  $B_{\text{unit}}^2 \gamma_{\text{max}}^2 / k_{\text{max}}^5$ , however for those that are TEM-dominated (cases f, x, y and z, shown in (b)), a scaling of  $B_{\text{unit}}^2 \gamma_{\text{max}}^2 \rho_{\text{unit}} / k_{\text{max}}^4$  is found to be in much better agreement, indicating a difference in saturation level between the two mode types.

To explain why this factor of  $k_{\text{max}} \rho_{\text{unit}}$  affects the isotope scaling, consider first this quantity in a system with adiabatic electrons. In such a situation, all quantities are exactly gyroBohm scaled, such that for three linear simulations in H, D and T the value of  $k_{\text{max}}$  occurs at the same value of  $k_y \rho_i$ . Because  $\rho_i \propto \sqrt{A}$ , one finds in the adiabatic electron case that  $k_{\text{max}} \propto A^{-0.5}$ . In the linear simulations of this work, for which kinetic electrons have been used, the non-adiabatic response causes the scaling of  $k_{\text{max}}$  with  $A$  to become slightly more positive at around  $\alpha_A \approx -0.3$  for both mode types, a factor which is sufficient to capture the differences between the isotope scalings of the two mode types' saturation levels seen in figure 5.8.

The above nonlinear kinetic electron physics will be captured in the saturation model by using two different saturation levels for the ITG and TEM, such that one has  $12.7 B_{\text{unit}}^2 \gamma_{\text{max}}^2 \rho_{\text{unit}} / k_{\text{max}}^4$

for the TEM, and  $3.3B_{\text{unit}}^2\gamma_{\text{max}}^2/k_{\text{max}}^5$  for the ITG, where the proportionality constants are taken from figure 5.8. In order to decide which saturation level to use for a given simulation, the linear physics must be considered to reveal whether the turbulence is ITG- or TEM-dominated. For SAT3, the ratio between the magnitude of the linear energy weights of the electrons and ions is used,  $|W_{e,k_y}^L|/|W_{i,k_y}^L| = |W_{e,k_y}^L/W_{i,k_y}^L|$ , due to the disparate behaviour in the ratio of the species' energy fluxes for the two mode types [106]. For ion-dominated turbulence, in which the electron energy fluxes are comparatively small, one expects the ratio of  $|W_{e,k_y}^L/W_{i,k_y}^L|$  to also be small. When in the regime of TEM turbulence however, the electron turbulent energy flux increases to approximately the level of the ions. The ratio of  $|W_{e,k_y}^L/W_{i,k_y}^L|$  can be calculated from linear physics, and so represents a useful metric to select between the two saturation levels. If new classes of modes were present, this aspect of the saturation rule would likely need to be revisited and extended further.

A plot of  $|W_{e,k_y}^L/W_{i,k_y}^L|$  for all cases in the database against  $k_y/k_{\text{max}}$  is shown in figure 5.9, with ITG cases in red and TEM cases in blue. It can be seen that the TEM cases cluster around a value of  $|W_{e,k_y}^L/W_{i,k_y}^L| \approx 1$ , whereas ITG cases are mainly grouped around 0.4.

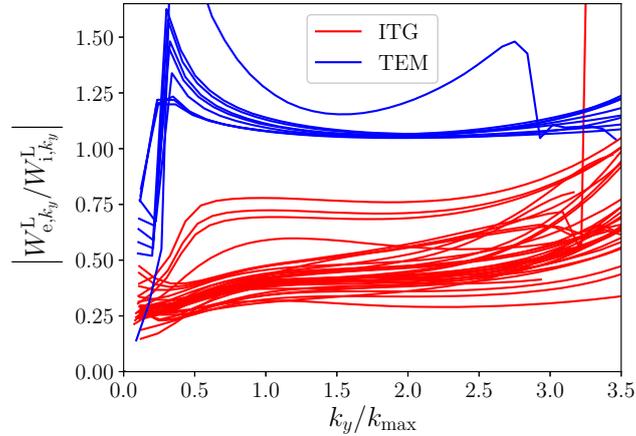


Figure 5.9: Ratios of the linear electron energy weight magnitudes to the linear ion energy weight magnitudes  $|W_{e,k_y}^L/W_{i,k_y}^L|$  against  $k_y/k_{\text{max}}$  for the database. Cases for which the dominant linear instability is the TEM are in blue, and those with ITG are shown in red. Note the disparate grouping between the mode types.

A transition function is now defined for the two modes, as a function of the weight ratio evaluated at  $k_y = k_{\text{max}}$ . At values below  $|W_{e,k_y}^L/W_{i,k_y}^L|_{k_y=k_{\text{max}}} = 0.8$  the ITG scaling is used, and at values of 1 and above the TEM scaling is used, with a first-order polynomial in-between to connect the two regions. The mode transition function  $M(x; x_1, x_2, y_1, y_2)$  is introduced

such that

$$M(x; x_1, x_2, y_1, y_2) = \begin{cases} y_1 & x \leq x_1 \\ y_1 \frac{x_2 - x}{x_2 - x_1} + y_2 \frac{x - x_1}{x_2 - x_1} & x_1 < x \leq x_2 \\ y_2 & x_2 < x \end{cases} \quad (5.17)$$

which allows the saturation level to be written

$$\frac{\langle |\delta \hat{\phi}_{k_y=k_0}|^2 \rangle_{x,\theta,t}}{\Delta k_y} = M \left( \left| \frac{W_{e,k_y}^L}{W_{i,k_y}^L} \right|_{k_y=k_{\max}} ; 0.8, 1.0, 3.3 \frac{\gamma_{\max}^2}{k_{\max}^5}, 12.7 \frac{\gamma_{\max}^2 \rho_{\text{unit}}}{k_{\max}^4} \right) B_{\text{unit}}^2. \quad (5.18)$$

It can be shown that measuring the sign of the linear frequency at a position in  $k_y$  can also be used to differentiate between the two mode types, however the above was chosen to attempt to avoid discontinuous changes in flux at the point of a mode transition.

## 5.5 Model extensions

Several effects have not been considered in the database of simulations that built SAT3, which will now be included. Some of these will come from novel considerations, while some can be incorporated from previous saturation rules.

### 5.5.1 Subdominant modes

The subdominant behaviour of previous TGLF saturation rules is taken forward into SAT3. Using equation 4.14, combined with the subdominant weight  $W_{s,k_y,m}^L$ , gives fluxes of the form

$$Q_s = 2 \sum_m \sum_{k_y > 0} \Lambda_{s,k_y} W_{s,k_y,m}^L \left[ \frac{\langle |\delta \hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t}}{\Delta k_y} \right] \left( \frac{\gamma_{k_y,m}}{\gamma_{k_y,m=1}} \right)^4 \Delta k_y \quad (5.19)$$

where the  $m = 1$  label for the saturated potentials of the dominant mode has been dropped for convenience.

### 5.5.2 $E \times B$ shear

Although no cases with  $E \times B$  shear were considered in the nonlinear database, one can simply isolate the  $E \times B$  effect of a previous rule and incorporate it into the new model. For SAT3 we choose to use that defined in SAT2, equation 4.18. Note this is not the only choice for  $F_{s,k_y}$ , one could for example have implemented a 1D quench rule. The effects of different  $E \times B$  rules can be tested using this freedom.

### 5.5.3 Mixed plasmas

For all simulations in the database of this work a pure plasma was used. For generality, quasilinear models must be able to operate with multiple ion species. The only aspect of SAT3 that is explicitly affected by this generalisation is the ratio of the magnitude of the linear weights  $\left|W_{e,k_y}^L/W_{i,k_y}^L\right|$ , used as an argument in the mode transition function, as this becomes ambiguously defined in a plasma with multiple ion species.

This ambiguity is resolved by changing the denominator of the linear weight ratio to the sum over ion weights,  $\left|W_{e,k_y}^L/W_{i,k_y}^L\right| \rightarrow \left|W_{e,k_y}^L/\sum_i W_{i,k_y}^L\right|$  for ion species  $i$ . By conducting a scan over relative concentrations of D and T in the GA-std and  $a/L_n = 3.0$  cases linearly (figure 5.10), this new ratio is shown to be invariant with relative density and thus recreate the expected behaviour.

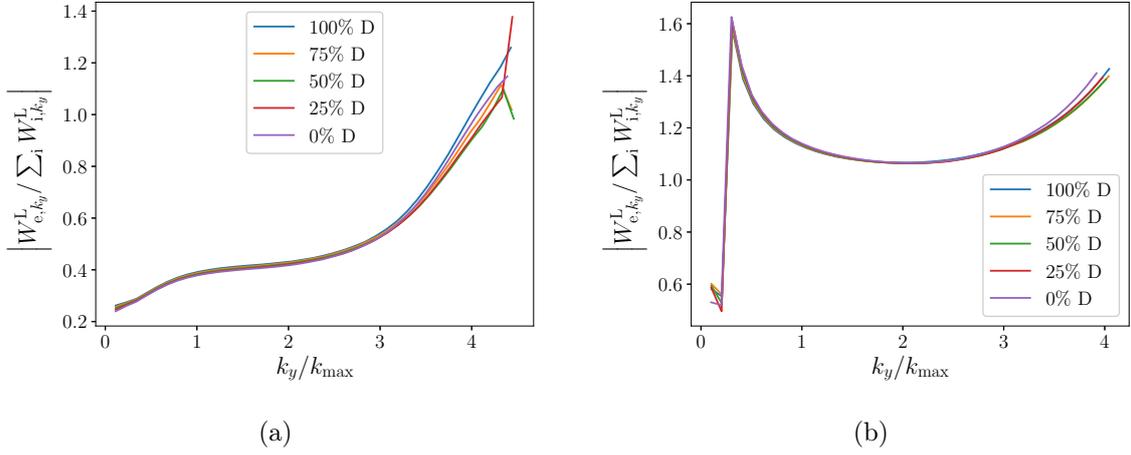


Figure 5.10: Linear energy weight magnitude ratios  $\left|W_{e,k_y}^L/\sum_i W_{i,k_y}^L\right|$  for different relative concentrations of D and T against  $k_y/k_{\max}$  for the GA-std case (a) and the  $a/L_n = 3.0$  case (b). Note the essential invariance across the spectrum for the different concentrations.

### 5.5.4 Electron scale

All simulations in the database considered simulated only the ion scale, up to  $k_y\rho_i = 1.0$ , where  $\rho_i = \sqrt{m_i/m_D}\rho_{\text{unit}}$ . Multiscale simulations [119, 136] are extremely computationally expensive, as one is required to simulate wavenumbers from the ion scale up to those comparable to the electron gyroradius,  $k_y\rho_e \sim 1$ , and thus  $k_y\rho_i \sim \sqrt{m_i/m_e} \sim 43\sqrt{A} \gg 1.0$ . Their expense renders them limited in abundance, however previous studies have been conducted, and electron scale rules informed by multiscale simulations have been implemented in SAT1, SAT2 and QuaLiKiz. In a similar manner as before, we may take a rule described previously

and, using the flexibility of our approach, incorporate it into SAT3.

We start by specifying a wavenumber,  $k_T$ , beyond which we declare our potentials to be electron scale. For  $k_y \leq k_T$ , we use the SAT3 model derived thus far, which we denote using a prime, with said electron-scale rule applying for  $k_y > k_T$ . In this work we use that of SAT2. To greatly reduce notational clutter, we introduce the variable  $Y_{\text{SAT2}}(k_y)$  to describe the saturated potentials of SAT2, such that from equation 4.17 we have

$$Y_{\text{SAT2}}(k_y) = \frac{\langle G^2(\theta) \rangle_\theta}{k_{\text{max}}} \left( \frac{\gamma^{\text{eff}}}{k_x^{\text{model}} k_y} \right)^2 \quad (5.20)$$

where these are the potentials of the dominant mode without  $E \times B$  shear, as in SAT3 both of these effects are included in the final flux summation. We also label the saturation level of SAT3, given by equation 5.18, using  $Y_0$  such that

$$Y_0 = M \left( \left| \frac{W_{e,k_y}^L}{W_{i,k_y}^L} \right|_{k_y=k_{\text{max}}} ; 0.8, 1.0, 3.3 \frac{\gamma_{\text{max}}^2}{k_{\text{max}}^5}, 12.7 \frac{\gamma_{\text{max}}^2 \rho_{\text{unit}}}{k_{\text{max}}^4} \right) B_{\text{unit}}^2. \quad (5.21)$$

The potentials for SAT3 are therefore written

$$\frac{\langle |\hat{\delta\phi}_{k_y}|^2 \rangle_{x,\theta,t}}{\Delta k_y} = \begin{cases} Y_0 \left( \frac{\sigma_{k_y}}{\sigma_{k_y=k_0}} \right)^{c_1} & 0 < k_y \leq k_T \\ Y_{\text{SAT2}}(k_y) & k_T < k_y < \infty. \end{cases} \quad (5.22)$$

Currently, these two branches do not connect, as in general  $Y_0 (\sigma_{k_y=k_T}/\sigma_{k_y=k_0})^{c_1} \neq Y_{\text{SAT2}}(k_y = k_T)$ .

This can be achieved however simply by making the linear part of the spectral width model (equation 5.12) into a quadratic,  $(2ak_P + b)k_y + c - ak_P^2 \rightarrow dk_y^2 + ek_y + f$ , where  $d, e$  and  $f$  are to be determined, and the domain is now  $k_P < k_y \leq k_T$ . The new model for the radial widths is therefore

$$\sigma_{k_y} = \begin{cases} ak_y^2 + bk_y + c & 0 < k_y \leq k_P \\ dk_y^2 + ek_y + f & k_P < k_y \leq k_T. \end{cases} \quad (5.23)$$

We first specify  $k_T$  to be the position at which  $k_y \rho_i = 1.0$ , and so  $k_T \rho_{\text{unit}} = \rho_{\text{unit}}/\rho_i$ . To specify  $d, e$  and  $f$ , we enforce continuity of function and of gradient at  $k_P$ , as before, however we now enforce continuity of function at  $k_T$ , such that  $\sigma_{k_y=k_T} = \sigma_{k_y=k_0} (Y_{\text{SAT2}}(k_y = k_T)/Y_0)^{1/c_1}$ . We obtain the set of linear equations

$$dk_P^2 + ek_P + f = ak_P^2 + bk_P + c \quad (5.24)$$

$$2dk_P + e = 2ak_P + b \quad (5.25)$$

$$dk_T^2 + ek_T + f = g \quad (5.26)$$

where  $g = \sigma_{k_y=k_0} (Y_{\text{SAT2}}(k_y = k_T) / Y_0)^{1/c_1}$ . These can be inverted to obtain

$$d = \frac{g + (k_P - k_T)(2ak_P + b) - (ak_P^2 + bk_P + c)}{(k_P - k_T)^2} \quad (5.27)$$

$$e = 2k_P(a - d) + b \quad (5.28)$$

$$f = g - dk_T^2 - (2k_P(a - d) + b)k_T \quad (5.29)$$

where  $e$  and  $f$  have been written in terms of  $d$  to reduce clutter.

The new normalised radial spectral widths are therefore

$$\frac{\sigma_{k_y}}{\sigma_{k_y=k_0}} = \begin{cases} \left( \frac{\frac{a}{b}k_y^2 + k_y + \frac{c}{b}}{\frac{a}{b}k_0^2 + k_0 + \frac{c}{b}} \right) & 0 < k_y \leq k_P \\ \left( \frac{d}{\sigma_{k_y=k_0}} \right) k_y^2 + \left( \frac{e}{\sigma_{k_y=k_0}} \right) k_y + \left( \frac{f}{\sigma_{k_y=k_0}} \right) & k_P < k_y \leq k_T \end{cases} \quad (5.30)$$

where

$$\frac{d}{\sigma_{k_y=k_0}} = \frac{1}{(k_P - k_T)^2} \left[ \left( \frac{Y_{\text{SAT2}}(k_y = k_T)}{Y_0} \right)^{1/c_1} + \frac{(k_P - k_T)(2\frac{a}{b}k_P + 1) - (\frac{a}{b}k_P^2 + k_P + \frac{c}{b})}{\frac{a}{b}k_0^2 + k_0 + \frac{c}{b}} \right] \quad (5.31)$$

$$\frac{e}{\sigma_{k_y=k_0}} = -2k_P \left( \frac{d}{\sigma_{k_y=k_0}} \right) + \frac{2\frac{a}{b}k_P + 1}{\frac{a}{b}k_0^2 + k_0 + \frac{c}{b}} \quad (5.32)$$

$$\frac{f}{\sigma_{k_y=k_0}} = \left( \frac{Y_{\text{SAT2}}(k_y = k_T)}{Y_0} \right)^{1/c_1} - \left( \frac{e}{\sigma_{k_y=k_0}} \right) k_T - \left( \frac{d}{\sigma_{k_y=k_0}} \right) k_T^2 \quad (5.33)$$

While complex, this is simply a quadratic over the range  $k_P < k_y \leq k_T$  which satisfies the boundary conditions

$$\begin{aligned} \frac{\sigma_{k_y=k_T}}{\sigma_{k_y=k_0}} &= \left( \frac{Y_{\text{SAT2}}(k_y = k_T)}{Y_0} \right)^{1/c_1} \\ \frac{\sigma_{k_y=k_P}}{\sigma_{k_y=k_0}} &= \frac{\frac{a}{b}k_P^2 + k_P + \frac{c}{b}}{\frac{a}{b}k_0^2 + k_0 + \frac{c}{b}} \\ \frac{1}{\sigma_{k_y=k_0}} \left. \frac{d\sigma_{k_y}}{dk_y} \right|_{k_y=k_P} &= \frac{2\frac{a}{b}k_P + 1}{\frac{a}{b}k_0^2 + k_0 + \frac{c}{b}} \end{aligned} \quad (5.34)$$

and so now we have connected the ion scale of SAT3 to an electron-scale rule. Because the electron-scale potentials of SAT2 are recreated entirely, the multiscale effect present in the SAT2 potentials are retained in SAT3.

## 5.6 Quasilinear approximation functions

Having described the saturated potentials, the QLA functions are now considered, defined in equation 4.6. In general for electrostatic turbulence there are  $3n_s$  of these functions, where  $n_s$  is the total number of species present, with one existing for each combination of the 3 velocity moments and species. Historically these functions have seen a comparatively small

amount of focus compared to the potentials, and are typically modelled as a constant [130]. By plotting these functions explicitly from the CGYRO data, one can determine to what degree a constant QLA is a reasonable assumption to make.

All cases in this database use kinetic electrons and a single ion species, giving 6 functions. Of these 6, the two momentum QLA functions are arbitrary due to there being no momentum transport in the cases considered<sup>7</sup>, and the two particle functions are identical as a consequence of ambipolarity. This leaves three as non-trivial: the two functions for the ion and electron heat transport, and one for the particle flux.

Plots of these three QLA functions against  $k_y \rho_i$  are shown in figure 5.11 for all cases in the database. The vast majority of the energy functions for both species have a similar shape across the spectrum, and most importantly exhibit relatively small variation in the region where the flux components are largest ( $k_y \rho_i \sim 0.2$ , evidenced in figure 4.4a). A similar description is seen for the particle function, although with more sporadic variation in some cases.

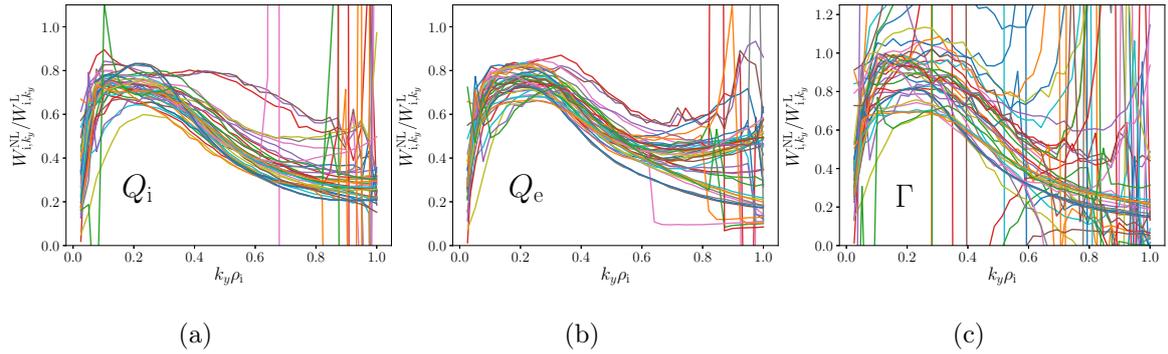


Figure 5.11: QLA functions for all cases in the database, showing those for (a) ion energy flux, (b) electron energy flux and (c) particle flux.

The assumption that the QLA could be modelled as constant was found to be a reasonable approximation for the database, given that the higher  $k_y$  part of the spectrum does not contribute significantly to the flux, with the constants for the model being set by minimising the scatter between the NL CGYRO flux data and the CGYRO-SAT3 flux data. The values of these constants for the electron and ion fluxes were found to be similar, however were stratified by mode type and moment. These are  $\Lambda_{ITG}^\Gamma = 1.1$ ,  $\Lambda_{TEM}^\Gamma = \Lambda_{TEM}^{Q_{i,e}} = 0.6$  and  $\Lambda_{ITG}^{Q_{i,e}} = 0.75$ . To capture these differences, the QLA functions are expressed in terms of the mode transition

<sup>7</sup>These are assumed below to be  $\Lambda_{s,k_y} = 0.8$ .

function (equation 5.17), such that for the particle flux

$$\Lambda_{s,k_y}^\Gamma = M \left( \left| \frac{W_{e,k_y}^L}{\sum_i W_{i,k_y}^L} \right|_{k_y=k_{\max}} ; 0.8, 1.0, 1.1, 0.6 \right) \quad (5.35)$$

and for the ion and electron energy fluxes

$$\Lambda_{s,k_y}^{Q_{i,e}} = M \left( \left| \frac{W_{e,k_y}^L}{\sum_i W_{i,k_y}^L} \right|_{k_y=k_{\max}} ; 0.8, 1.0, 0.75, 0.6 \right). \quad (5.36)$$

Any other forms of transport are assumed to have  $\Lambda_{s,k_y} = 0.8$ .

While a constant that depends on the dominant mode type of ITG or TEM remains a reasonable model for the contribution of the QLA functions to the overall flux in this database, it is not perfect, missing for example the isotope scaling seen in figure 4.5b. In future studies more exotic turbulence regimes may cause the QLA functions to deviate further from those seen here, potentially necessitating an effort to try to build a more sophisticated quasilinear description of phase relationships in the saturated state.

## 5.7 Summary of new saturation model, SAT3

The entirety of SAT3 is collected here for reference<sup>8</sup>. The fluxes of the model are constructed via

$$Q_s = 2 \sum_m \sum_{k_y > 0} \Lambda_{s,k_y} W_{s,k_y,m}^L F_{s,k_y} \left[ \frac{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}}{\Delta k_y} \right] \left( \frac{\gamma_{k_y,m}}{\gamma_{k_y,m=1}} \right)^4 \Delta k_y \quad (5.37)$$

where  $m$  is the mode number,  $\Lambda_{s,k_y}$  is the quasilinear approximation (QLA) function,  $W_{s,k_y,m}^L$  is the linear weight for species  $s$ ,  $F_{s,k_y}$  describes the effect of  $E \times B$  shear,  $\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t} / \Delta k_y$  is the saturated potential of the dominant mode, and  $\Delta k_y$  is the binormal grid spacing of the simulation. The model for the saturated potentials is

$$\frac{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}}{\Delta k_y} = \begin{cases} \frac{\left\langle \left| \delta \hat{\phi}_{k_y=k_0} \right|^2 \right\rangle_{x,\theta,t}}{\Delta k_y} \left( \frac{\sigma_{k_y}}{\sigma_{k_y=k_0}} \right)^{-2.42} & 0 < k_y \leq k_T \\ Y_{\text{SAT2}}(k_y) & k_T < k_y < \infty. \end{cases} \quad (5.38)$$

---

<sup>8</sup>This version of SAT3 is similar to that presented in [137], however now includes the effect of subdominant modes and electron-scale saturation.

where  $k_T = 1/\rho_i$  separates the ion and electron scales,  $Y_{\text{SAT2}}$  is the saturated potentials of SAT2 given by equation 5.20, and the saturation level is

$$\frac{\left\langle \left| \delta \hat{\phi}_{k_y=k_0} \right|^2 \right\rangle_{x,\theta,t}}{\Delta k_y} = M \left( \left| \frac{W_{e,k_y}^L}{\sum_i W_{i,k_y}^L} \right|_{k_y=k_{\max}} ; 0.8, 1.0, 3.3 \frac{\gamma_{\max}^2}{k_{\max}^5}, 12.7 \frac{\gamma_{\max}^2 \rho_{\text{unit}}}{k_{\max}^4} \right) B_{\text{unit}}^2 \quad (5.39)$$

where  $M$  is the mode transition function

$$M(x; x_1, x_2, y_1, y_2) = \begin{cases} y_1 & x \leq x_1 \\ y_1 \frac{x_2-x}{x_2-x_1} + y_2 \frac{x-x_1}{x_2-x_1} & x_1 < x \leq x_2 \\ y_2 & x_2 < x \end{cases} \quad (5.40)$$

which captures the disparate saturation levels between ITG- and TEM-dominated turbulence. The wavenumber  $k_{\max}$  is defined as the position of the maximum of the linear growth rate divided by  $k_y$ ,

$$\left. \frac{d}{dk_y} \left( \frac{\gamma_{k_y}}{k_y} \right) \right|_{k_y=k_{\max}} = 0 \quad (5.41)$$

and  $\gamma_{\max} = \gamma_{k_y=k_{\max}}$ . The model for the radial spectral widths is

$$\frac{\sigma_{k_y}}{\sigma_{k_y=k_0}} = \begin{cases} \left( \frac{a}{b} k_y^2 + k_y + \frac{c}{b} \right) / \left( \frac{a}{b} k_0^2 + k_0 + \frac{c}{b} \right) & 0 < k_y \leq k_P \\ \left( \frac{d}{\sigma_{k_y=k_0}} \right) k_y^2 + \left( \frac{e}{\sigma_{k_y=k_0}} \right) k_y + \left( \frac{f}{\sigma_{k_y=k_0}} \right) & k_P < k_y \leq k_T \end{cases} \quad (5.42)$$

where

$$\frac{d}{\sigma_{k_y=k_0}} = \frac{1}{(k_P - k_T)^2} \left[ \left( \frac{Y_{\text{SAT2}}(k_y = k_T)}{Y_0} \right)^{-1/2.42} + \frac{(k_P - k_T) \left( 2\frac{a}{b} k_P + 1 \right) - \left( \frac{a}{b} k_P^2 + k_P + \frac{c}{b} \right)}{\frac{a}{b} k_0^2 + k_0 + \frac{c}{b}} \right] \quad (5.43)$$

$$\frac{e}{\sigma_{k_y=k_0}} = -2k_P \left( \frac{d}{\sigma_{k_y=k_0}} \right) + \frac{2\frac{a}{b} k_P + 1}{\frac{a}{b} k_0^2 + k_0 + \frac{c}{b}} \quad (5.44)$$

$$\frac{f}{\sigma_{k_y=k_0}} = \left( \frac{Y_{\text{SAT2}}(k_y = k_T)}{Y_0} \right)^{1/c_1} - \left( \frac{e}{\sigma_{k_y=k_0}} \right) k_T - \left( \frac{d}{\sigma_{k_y=k_0}} \right) k_T^2 \quad (5.45)$$

and  $c/b = -0.751k_{\max}$ ,  $k_{\min} = 0.685k_{\max}$ ,  $k_0 = 0.6k_{\min}$  and  $k_P = 2k_{\min}$ .

The quasilinear approximation functions  $\Lambda_{s,k_y}$  vary depending on velocity moment and mode type. For the particle flux and energy flux these are

$$\Lambda_{s,k_y}^\Gamma = M \left( \left| \frac{W_{e,k_y}^L}{\sum_i W_{i,k_y}^L} \right|_{k_y=k_{\max}} ; 0.8, 1.0, 1.1, 0.6 \right) \quad (5.46)$$

$$\Lambda_{s,k_y}^{Q_{i,e}} = M \left( \left| \frac{W_{e,k_y}^L}{\sum_i W_{i,k_y}^L} \right|_{k_y=k_{\max}} ; 0.8, 1.0, 0.75, 0.6 \right) \quad (5.47)$$

with other fluxes taking  $\Lambda_{s,k_y} = 0.8$ .

Finally, the function  $F_{s,k_y}$  describes the effect of  $E \times B$  shear and is given by [52]

$$F_{s,k_y}(\gamma_E) = \frac{\tilde{W}_{s,k_y,k_x=k_{x0}}^L / \tilde{W}_{s,k_y,k_x=0}^L}{\left(1 + \left(\alpha_x \frac{k_{x0}}{k_x^{\text{model}}}\right)^{\sigma_x}\right)^2 \left(1 + \left(\frac{k_{x0}}{k_x^{\text{model}}}\right)^2\right)^2} \quad (5.48)$$

where  $\tilde{W}_{s,k_y,k_x}^L$  is the quasilinear weight defined by TGLF, which is evaluated at a single  $k_x$  rather than a sum over  $k_x$ . The two constants are  $\alpha_x = 1.21$  and  $\sigma_x = 2$ , with  $k_{x0} = 0.32k_y (k_{\text{max}}/k_y)^{0.7} (\gamma_E/\gamma_{\text{max}})$  and

$$k_x^{\text{model}} = \begin{cases} \frac{0.76k_{\text{max}}}{\sqrt{g^{xx}(\theta=0)}} & k_y < 0.76k_{\text{max}} \\ \frac{0.76k_{\text{max}}}{\sqrt{g^{xx}(\theta=0)}} + \frac{1.22}{\sqrt{g^{xx}(\theta=0)}} \frac{B(\theta=0)}{B_{\text{unit}}} (k_y - 0.76k_{\text{max}}) & k_y \geq 0.76k_{\text{max}} \end{cases} \quad (5.49)$$

where  $g^{xx} = |\nabla x|^2$  and  $B$  is the equilibrium magnetic field magnitude.

## 5.8 Results

The scatter plots of the fluxes obtained from NL CGYRO against the results of CGYRO-SAT3 are shown in figure 5.12 for the ion energy fluxes, electron energy fluxes and particle fluxes. A metric for the quality of the model agreement can be calculated by taking the average percentage error,

$$\Sigma_{Q_{i,e}} = \frac{1}{N} \sum_{m=1}^N \left| \frac{Q_{\text{model},m} - Q_{\text{NL},m}}{Q_{\text{NL},m}} \right| \quad (5.50)$$

$$\Sigma_{\Gamma} = \frac{1}{N} \sum_{m=1}^N \left| \frac{\Gamma_{\text{model},m} - \Gamma_{\text{NL},m}}{\Gamma_{\text{NL},m}} \right| \quad (5.51)$$

for the energy fluxes and particle flux, where  $N$  is the number of simulations in the database. The values obtained for the three plots are displayed in their respective subfigures.

For comparison, the equivalent scatter plots for a rescaled CGYRO-SAT1 model are shown in figure 5.13, with additional fitted prefactors for each flux type, labelled CGYRO-SAT1\*<sup>9</sup>. Looking at figures 5.12 and 5.13 a reduction in the average percentage error between CGYRO-SAT1\* and CGYRO-SAT3 is seen for the three flux types. CGYRO-SAT1\* and CGYRO-SAT3 can be seen to perform similarly in the ITG-dominated cases, as may be expected, however

<sup>9</sup>This accounts for the larger NL CGYRO fluxes compared to those of SAT1's training database, which resulted from an increased binormal resolution in this work. These pre-factors were fitted to the D simulations of the ITG-dominated cases, and have values of 2.16 for the ion energy fluxes, 2.23 for the electron energy fluxes and 2.98 for the particle fluxes. Note that the isotope scaling of the fluxes is unaffected by a database-wide rescaling constant.

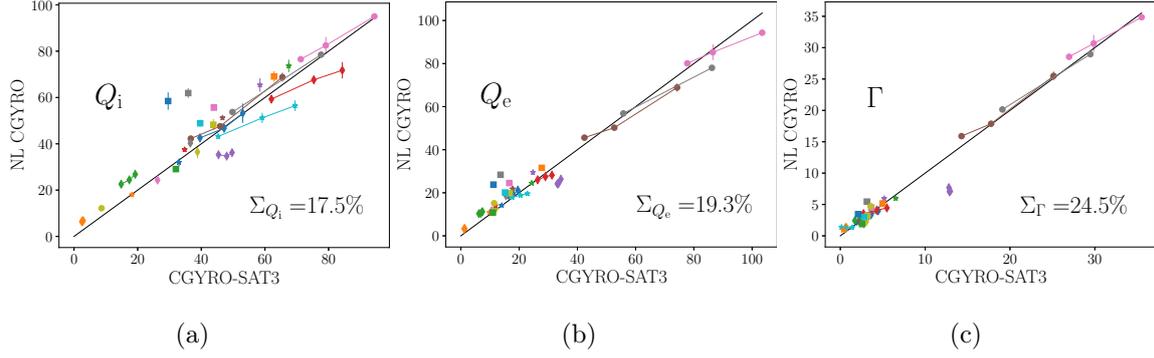


Figure 5.12: Scatter plots of NL CGYRO against CGYRO-SAT3 results for (a) ion energy flux, (b) electron energy flux and (c) particle flux. The legend for these figures is shown in figure 5.8c. TEM-dominated cases are marked by circles (labelled f, x, y and z in figure 5.8c), with the remainder being ITG-dominated. Data points connected by a line denote simulations of the same equilibrium but different of isotopes. The line to denote perfect agreement between the models is shown in black, with the respective relative errors also shown on each plot, defined by equations 5.50 and 5.51.

great improvement is shown in the isotope scaling and magnitude of the TEM cases, owing to the modelling of the difference in saturation level between the two mode types present in SAT3. This is demonstrated explicitly in figure 5.14, which exhibits the ion and electron energy fluxes against isotope mass for the ITG-dominated GA-std case (5.14a) and the TEM-dominated  $a/L_n = 3.0$  case (5.14b) compared between the three models.

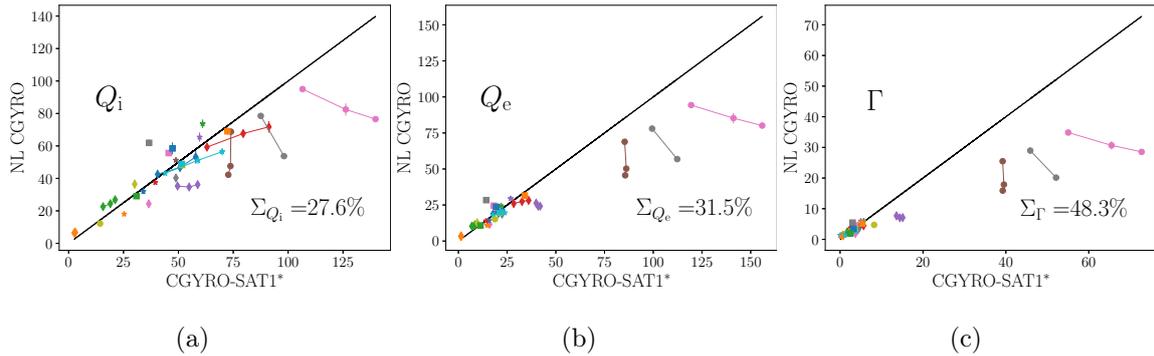


Figure 5.13: Scatter plots equivalent to those of figure 5.12, however now comparing NL CGYRO against CGYRO-SAT1\* results for (a) ion energy flux, (b) electron energy flux and (c) particle flux.

A selection of energy flux scans with various key tokamak parameters is shown in figure 5.15, comparing NL CGYRO and CGYRO-SAT3. The density gradient scan in 5.15a is

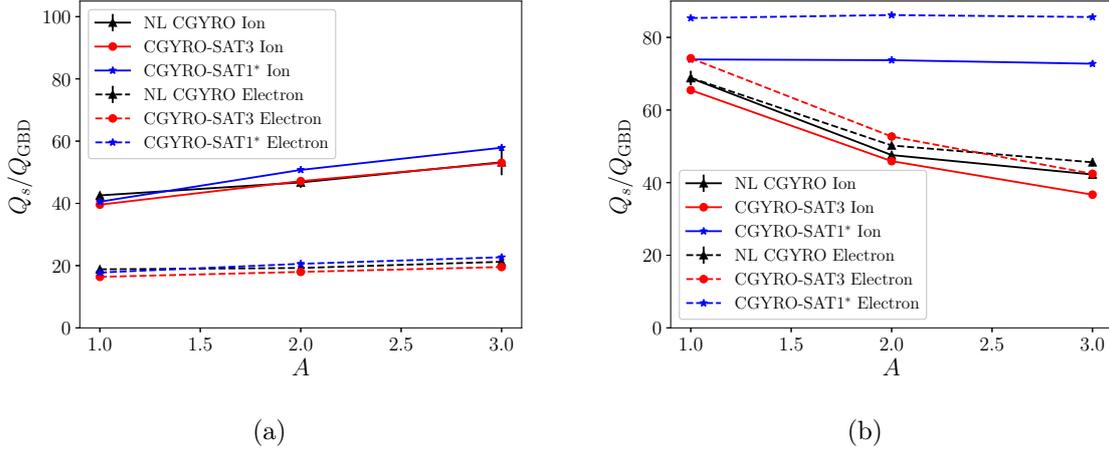


Figure 5.14: Turbulent ion energy fluxes (solid lines) and electron energy fluxes (dashed lines) against isotope mass for NL CGYRO (black), CGYRO-SAT3 (red) and CGYRO-SAT1\* (blue), for the ITG-dominated GA-standard case (a) and the TEM-dominated  $a/L_n = 3.0$  case (b). Note the recreation of the anti-gyroBohm scaling seen in the  $a/L_n = 3.0$  case for CGYRO-SAT3.

a recreation of a subset of the data points from figure 4.2, now with larger NL CGYRO fluxes, demonstrating the recreation of the positive isotope scaling for the ITG-dominated GA-std cases at low density gradient, the grouping of the fluxes for the transition case, and the anti-gyroBohm scaling at the high density gradient TEM-dominated  $a/L_n = 3.0$  case. How this  $a/L_n = 3.0$  case varies with collisionality is then displayed in figure 5.15b, from which it is seen that the correct anti-gyroBohm scaling is maintained across a large range of collisionalities. Finally the ion and electron heat fluxes against matched temperature gradients ( $a/L_{T_i} = a/L_{T_e}$ ) are shown in figure 5.15c. The general trend in both isotope scaling and magnitude can be observed to agree with the NL CGYRO data, however the fluxes appear to be somewhat under-predicted near the point of threshold. Note that this is also present in CGYRO-SAT1\*, observing that these low temperature gradient equilibria, labelled b and c in figure 5.8c (the orange and green diamonds respectively), are similarly underpredicted in figures 5.13a and 5.13b. The generality of SAT3's behaviour in this region presents an area to be investigated further, as this is a key region of parameter space for experimental conditions. To connect these results with the observations made in Section 4.7, figure 5.16a shows a recreation of figure 4.4b, exhibiting the isotope scaling of the ion energy flux components  $Q_{i,k_y}$  in the  $a/L_n = 3.0$  case for the different models, now with the data for CGYRO-SAT3 included. The improved low  $k_y \rho_i$  variation and offset in the isotope scaling can be seen, as a consequence of the strong spectral shape and the differing TEM saturation level in SAT3. In

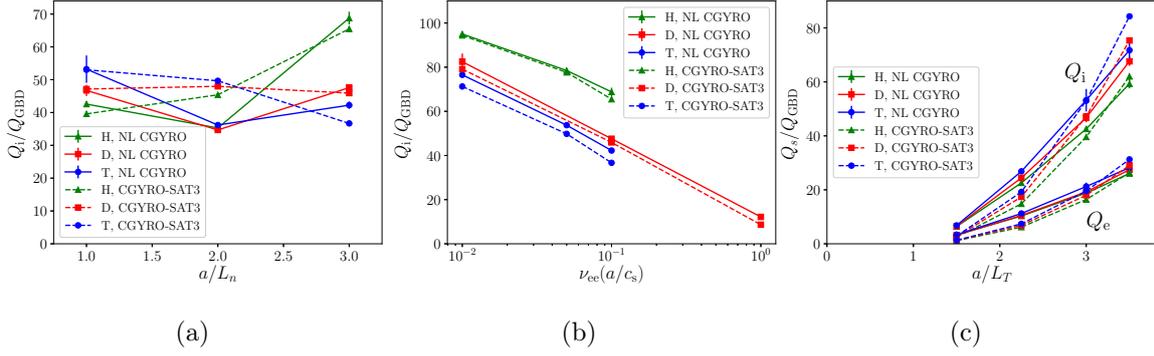


Figure 5.15: (a) Ion energy fluxes for NL CGYRO (solid lines) and CGYRO-SAT3 (dashed lines) against density gradient scale length  $a/L_n$ . (b) Ion energy fluxes for NL CGYRO and CGYRO-SAT3 against collisionality, keeping  $a/L_n = 3.0$  fixed. (c) Ion and electron energy fluxes for NL CGYRO and CGYRO-SAT3 against matched temperature gradient scale length,  $a/L_{T_i} = a/L_{T_e}$ .

figure 5.16b the isotope scaling of the saturated potentials in the GA-std case are presented for NL CGYRO and the reduced models. CGYRO-SAT3 is seen to maintain the relevant isotope scaling in this ITG-dominated case.

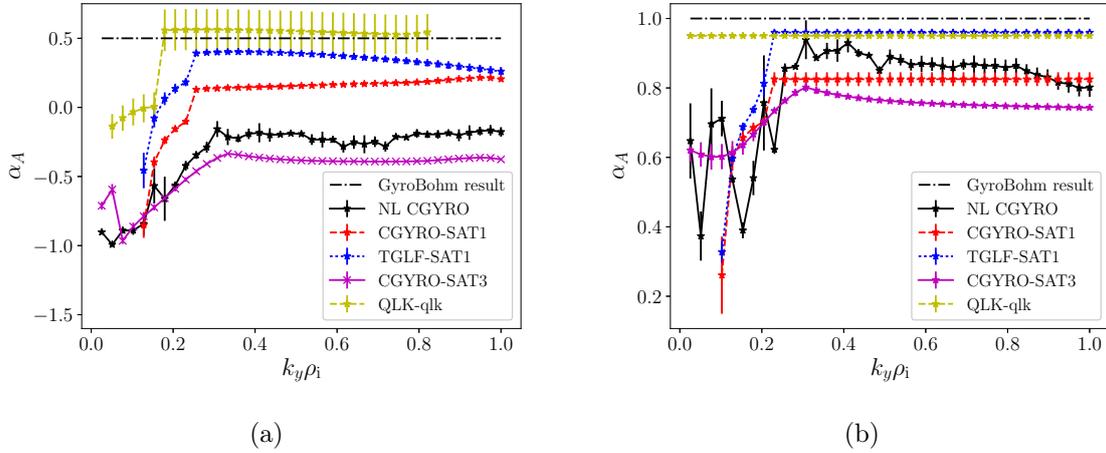


Figure 5.16: (a) Isotope scaling metric  $\alpha_A$  of the flux components for the  $a/L_n = 3.0$  case, for NL CGYRO (black), CGYRO-SAT3 (purple), CGYRO-SAT1 (red), TGLF-SAT1 (blue) and QuaLiKiz-qlk (yellow). (b) The isotope scaling of the saturated potentials in the GA-std case for the same models. The error bars shown for all models are the uncertainties in the fitted parameter  $\alpha_A$ .

## 5.9 A note on future improvements

The SAT3 model presented in this chapter is an accurate model for the enhanced CGYRO database generated, and is promising for extrapolation within the most common ITG and TEM turbulence regimes found in experiment. It does not however represent a complete description of turbulent plasma transport, and will of course perform less reliably in a parameter space far from its training database. The approach taken in SAT3's development however naturally presents a methodology with which to diagnose and improve on such future discrepancies. That is, for some hypothetical simulation that this model catastrophically fails to recreate, the methodology established in this chapter can be employed to effectively diagnose and attempt to rectify the issue.

This approach starts by taking care to preserve the role of each physical element that constitutes the model flux calculation, so as to keep separate the contributions of the QLA, the weights, and the potentials. This decomposition of the reduced model (equation 4.7) clarifies which aspect of the reduced model is responsible for the recreation of each part of the non-linear flux, and each can be considered in turn when attempting to diagnose future model disagreements with NL gyrokinetic results.

If the discrepancy is found to originate in the saturation rule, then one can first test whether the underlying functional relations (equations 5.38 and 5.42) still hold when their parameters are fitted to nonlinear data. If not, these will need to be amended. Otherwise, one continues to the question of the linear modelling of the parameters  $c_1$ ,  $k_{\min}$ ,  $c/b$  and  $\left\langle \left| \delta \hat{\phi}_{k_y=k_0} \right|^2 \right\rangle_{x,\theta,t} / \Delta k_y$ , each of which can be considered modularly. If a more robust method for approximating  $k_{\min}$  from the linear data is discovered in the future for example, one can replace/amend that aspect of the model without requiring change elsewhere, forming incremental and guided improvements.

## Chapter 6

# Validation against JET isotope experiments

### 6.1 Introduction

The verification of the new saturation rule SAT3 for the training database was shown at the end of the previous chapter, in which it was seen that the model is able to recreate the anti-gyroBohm scaling observed in the nonlinear gyrokinetic simulations considered. This represents a step forward for the comparison of quasilinear models to nonlinear gyrokinetics in stand-alone, however this new saturation rule must now be validated against experimental data. In this chapter we validate the new saturation rule SAT3 in predictive integrated modelling simulations using unique sets of 3-isotope experiments in JET.

During the course of this PhD, JET carried out its second D-T campaign, DTE-2, performing plasma pulses in H, D, and T as well as mixtures of these isotopes, including H-T and D-T. This was done to further probe and understand the experimental isotope effect, in part to help inform on the future performance of ITER. The diagnostics used during this campaign represent the state-of-the-art [138], with the experimental measurements generally being of high quality. Access to this data, combined with that from D and H experiments from previous campaigns, provides the unique opportunity to validate the performance of SAT3 in integrated modelling against 3 different isotopes, as well as make comparisons to other contemporary quasilinear models. All experiments considered in this work were carried out after the installation of JET's ITER-like-wall (ILW). This replaced the previous components of the tokamak chamber with plasma facing materials representing those chosen for the D-T operations of ITER, namely an all-metal combination of beryllium and tungsten, motivated

in part by lower amounts of tritium retention [139].

The new model SAT3 was paired with the linear solver of TGLF for use in the integrated modelling suite JETTO. Whilst a degree of error is incurred in flux prediction by moving from an exact linear solver to a fast linear solver, for most cases (but not all) this effect is predicted to be small, with fast linear solvers having shown good agreement with linear gyrokinetics in the most common experimentally-relevant regimes [50].

## 6.2 JETTO

All integrated modelling simulations carried out in this work were performed with the code JETTO [32]. JETTO couples together several models describing different aspects of tokamak operation to solve the radial transport equations for each plasma species, as described in section 2.5.

The transport equations that are solved include the particle continuity equations for each individual ion species, the energy equations for the electrons and the total ions, the total ion toroidal momentum equation<sup>1</sup>, and the evolution of the equilibrium magnetic field via Faraday’s law. This last equation is sometimes described as evolving the current or the safety factor  $q$ , due to their interconnection illustrated in section 3.4. The transport equations are evolved in time as a function of minor radius.

As shown schematically by the terms of equation 2.18, several features inform the evolution of the macroscopic plasma profiles. These include sources and sinks (e.g.  $S_{ns}$ ), such as those associated with auxiliary heating mechanisms, radiation losses and fusion reactions, the calculation of fluxes ( $\Gamma_s$ ), including classical, neoclassical and turbulent transport, and the magnetic equilibrium geometry ( $V'$ ), including the effect of MHD processes. Models are present within JETTO to simulate all such effects.

Those most pertinent to this work are the models for local turbulent transport, which is performed by quasilinear models. For a grid spanning the minor radius, a local equilibrium is defined at each point, informed by the experimental radial profiles and geometry. Each call of the transport solver calculates the turbulent fluxes at each grid point, which is then used in the radial transport equations. Even given their relative speed in computing fluxes as compared with nonlinear gyrokinetics, quasilinear models are the most computationally-costly aspect of JETTO simulations, accounting for the overwhelming majority of the runtime. Analytical

---

<sup>1</sup>Only having a single ion energy equation and a single ion momentum equation comes from the assumption that all ion species have the same temperature and toroidal rotation.

models for classical transport are also present, as well as models for neoclassical transport such as NCLASS [140]. The only source of auxiliary heating used in the discharges analysed in this chapter is NBI, which is simulated self-consistently using the code PENCIL [141].

If the sources and boundary conditions are constant in time, including those obtained by averaging over an experimental window, then the transport equations will be evolved until the system ‘relaxes’ to an equilibrium. As initial profiles are taken from the experimental data, radial profiles that are minimally changed over the simulation indicate a well-modelled discharge. It is generally seen however that the profiles first vary before settling to a new equilibrium, predominantly due to differences in levels of predicted transport<sup>2</sup>.

For a given set of input radial profiles, in JETTO one can choose to simulate each transport channel (ion density, electron temperature, etc) either *predictively*, in which it is evolved via the radial transport equations, or *interpretively*, for which the profile is unchanged from its experimental input. A simulation in which all profiles are simulated interpretively therefore represents the experimental plasma, with which to compare to predictive simulations. In this work the current is always modelled predictively, since it is not routinely measured accurately. Toroidal rotation is generally modelled interpretively, due to the difficulties in its prediction from quasilinear turbulence models. In this work, ‘interpretive’ simulations correspond to those in which all channels other than current are simulated interpretively, and ‘predictive’ simulations model the evolution of the current, ion temperature and electron temperature predictively. The density and toroidal rotation are simulated interpretively in all cases. The specifics of the discharges analysed in this work will now be detailed.

## 6.3 JET pulses

### 6.3.1 L-mode pulses

Three L-mode discharges are analysed, one in each hydrogenic isotope. These are #91450 (H), #89723 (D) and #99173 (T), all over a time window of 55 – 56 s during which NBI heating was used. Several key experimental measurements of these discharges are presented in figure 6.1. These JET-ILW experiments were conducted with a current of 2.5 MA and a magnetic field strength of 3.0 T, and were designed to operate at a similar line-averaged<sup>3</sup> density<sup>4</sup> of  $\bar{n}_e \sim 3.0 \times 10^{19} \text{ m}^{-3}$  and total stored energy, as shown in the lower panels of

<sup>2</sup>Errors associated with profile measurements can also be an influence here.

<sup>3</sup>Averaged across the line of sight through the plasma.

<sup>4</sup>The oscillations seen in the tritium line-averaged density correspond with  $n = 1$  mode MHD activity, the associated core transport of which is incorporated into the simulations of this chapter via a continuous

figures 6.1a and 6.1b respectively<sup>5</sup>. To achieve this, disparate levels of NBI heating and gas puffing rates were required between the shots, shown in the top panels of figures 6.1a and 6.1b, to account for the variation of confinement of energy and particles with isotope. The H and D experiments were found to have a positive (anti-GyroBohm) isotope scaling of the confinement time,  $\tau_{E,th} \sim A^{0.15 \pm 0.02}$  [61], manifesting the isotope effect. The H discharge therefore required the most additional heating at 4.38 MW, with D requiring 3.20 MW and T 1.45 MW.

These cases are of particular interest due to the presence of TEM-dominated turbulence in the L-mode edge [41, 43], which SAT3's second saturation branch looks to capture in the turbulent fluxes.

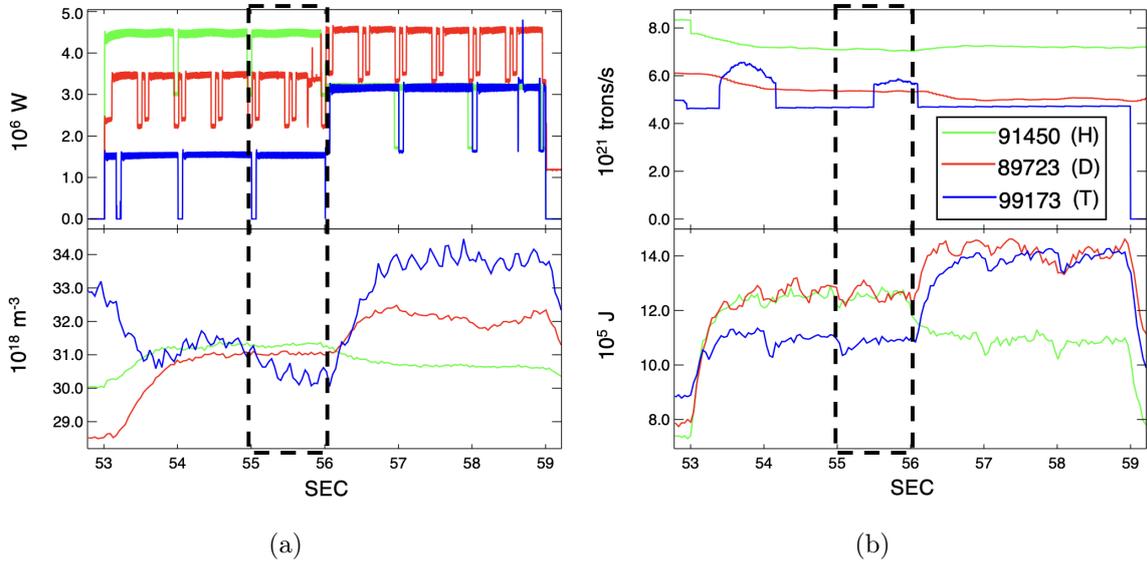


Figure 6.1: Experimental time-traces of the L-mode discharges, with the dashed boxes indicating the window analysed in this chapter. From top to bottom, shown in (a) are the NBI powers and line-averaged densities, and in (b) are the gas puffing rates and total stored energy. Note the differing levels of NBI and puffing required to achieve the similarity in the stored energies and densities between the pulses.

sawtooth model, as discussed in section 6.4.

<sup>5</sup>Pulses #91450 (H) and #89723 (D) were conducted as part of a D and H campaign in 2016, with #99173 (T) being performed as part of DTE-2 in 2021, and is thus not as well matched in terms of density and stored energy.

## Input profiles

In order to generate radial input profiles for JETTO, the experimental data of the electron density and electron temperature from HRTS (High Resolution Thompson Scattering) [142] was fitted in Profile Maker, with the profiles of angular frequency profiles being fitted to the best available diagnostic in operation during the pulse based on charge exchange spectroscopy measurements. Within experimental uncertainties,  $T_i = T_e$  is assumed for the analysis of this chapter [61]. The initial safety factor profiles were generated using EFIT [143].

### 6.3.2 Ohmic pulses

Twelve Ohmic pulses from a LOC-SOC density scan were analysed in this work, which are grouped into four sets of three. The three simulations in each set correspond to the three isotopes of H, D and T, and the four sets are differentiated by the magnitude of the line-averaged density, which affects their confinement properties within the Ohmic regime, as discussed in section 2.4. A plot of confinement time against line-averaged density is shown in figure 6.2 for all pulses in the density scan, with the twelve cases analysed circled in black. Note the approximately linear relation in the region of low density, corresponding to the LOC regime, and the flat region for the higher densities, which is the SOC regime. The LOC regime is of particular interest for the validation of SAT3 due to the presence of TEM-dominated turbulence, as demonstrated in previous studies [144].

Set 1 (#91633 (H), #96281 (D), #100145 (T),  $\bar{n}_e \sim 0.87 \times 10^{19} \text{ m}^{-3}$ ) are pulses of low-density, sitting in the LOC regime. Sets 3 (#91634a (H), #97553 (D), #100112a (T) ,  $\bar{n}_e \sim 2.30 \times 10^{19} \text{ m}^{-3}$ ) and 4 (#91634b (H), #95766 (D), #100112b (T),  $\bar{n}_e \sim 2.80 \times 10^{19} \text{ m}^{-3}$ ) are cases of high-density, which exhibit behaviour of the SOC regime, and set 2 (#91637, #90633, #99263,  $\bar{n}_e \sim 1.45 \times 10^{19} \text{ m}^{-3}$ ) is a mid-density set, lying in the transition between the regimes of confinement. The labels of ‘a’ and ‘b’ refer to different time windows for the same experimental discharge. All pulses were conducted with a current of 2.3 MA and a magnetic field strength of 2.7 T, and were analysed over a time window of 0.5 s. A summary of the relevant parameters for these discharges is shown in table 6.1.

Set	Pulse	$\bar{n}_e$ ( $10^{19}\text{m}^{-3}$ )	Time window start (s)	$Z_{\text{eff}}$ (TRANSP)	$T_e$ factor
1	91633 (H)	0.831838	53.4110	1.41	1.0
	96281 (D)	0.883970	59.0125	2.01	1.0
	100145 (T)	0.883229	53.1125	1.34	1.0
2	91637 (H)	1.47234	58.2510	1.05	1.15
	90633 (D)	1.44546	53.2110	1.41	1.125
	99263 (T)	1.43482	56.1125	1.05	1.075
3	91634a (H)	2.33095	55.9110	1.05	1.2
	97553 (D)	2.243581	56.2125	1.05	1.125
	100112a (T)	2.33841	56.1125	1.05	1.15
4	91634b (H)	2.83490	60.4110	1.05	1.175
	95766 (D)	2.81479	59.2125	1.37	1.15
	100112b (T)	2.74171	59.1125	1.05	1.125

Table 6.1: Experimental details of the 12 Ohmic cases simulated. Note that the measured  $Z_{\text{eff}}$  values are unreliable due to the difficulty in its experimental measurement. A subset of the values are hence changed for the simulations of this chapter, as discussed in the text.

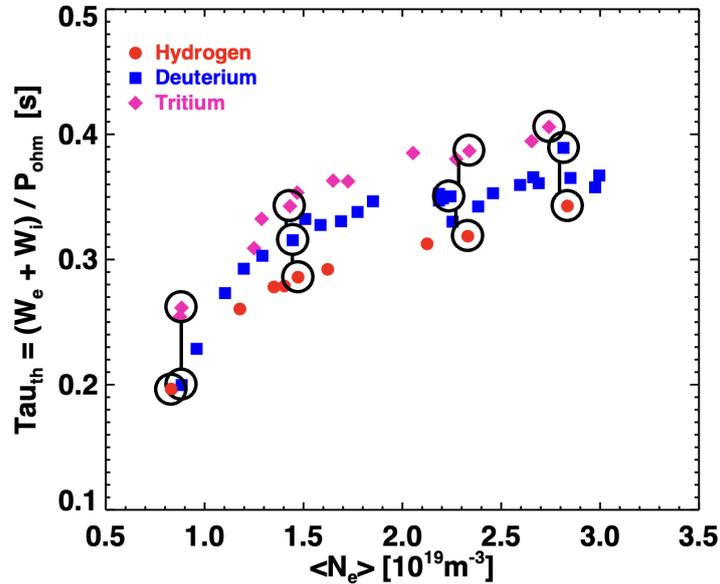


Figure 6.2: Confinement time against line-averaged density for cases in a LOC-SOC scan for H, D and T. The four sets are grouped by connected circles, with sets 1, 2, 3 and 4 appearing in order of increasing density. Figure adapted from [145].

## Input profiles

The profiles for the initial conditions of the radial profiles for JETTO were taken from TRANSP [146] runs, for which the electron temperature profiles was taken from HRTS, and the ion temperature profiles were obtained using main ion charge exchange measurements using beam blips [147, 148]. It was found however that the measured Ohmic power, a quantity of primary importance in Ohmic discharges, was not in agreement with the values measured by EFIT and EFTP, as shown in figure 6.3. This is due to their differing approaches in its calculation, with EFIT providing a more robust value, calculating it only from the loop voltage and plasma current, whereas TRANSP and JETTO use a resistivity model to compute local heating and therefore depend on the kinetic profiles and  $Z_{\text{eff}}$ <sup>6</sup>, making it more sensitive to errors in profile measurements and mappings.

The Ohmic power for the input profiles was therefore first brought into agreement with EFIT. The  $Z_{\text{eff}}$  measurements in JET are difficult to make, and thus have large error bars and pulse-to-pulse variations, as evidenced by the penultimate column of table 6.1. Low power discharges such as these Ohmic cases are expected to have consistently very clean plasma, and so we are justified in altering these values. The Ohmic power increases with  $Z_{\text{eff}}$ , and so these were reduced to a more representative value for all cases, at  $Z_{\text{eff}} = 1.05$ .

When simulated interpretively in JETTO, this exercise brought the Ohmic power into agreement with EFIT for the low-density set, (cases 1-3 in figure 6.3). For the remaining sets, the trend of the Ohmic power with case was now resembled more closely, however there was still disagreement in the magnitude. Due to the mechanism of Ohmic heating originating from ion-electron collisions, which have a strong dependence on the electron temperature  $\propto T_e^{-3/2}$ , the flexibility in the uncertainty of the experimental profiles was used to bring the Ohmic powers into agreement by multiplying the electron temperature profile by the necessary factor. The required factors varied with case, and are presented in the rightmost column of table 6.1.

---

<sup>6</sup>‘Z effective’, a quantity related to the average ion charge in the plasma that parameterises the radiative losses due to Bremsstrahlung. For pure hydrogenic plasmas  $Z_{\text{eff}} = 1.0$ , with the presence of  $Z \neq 1$  impurities increasing its value.

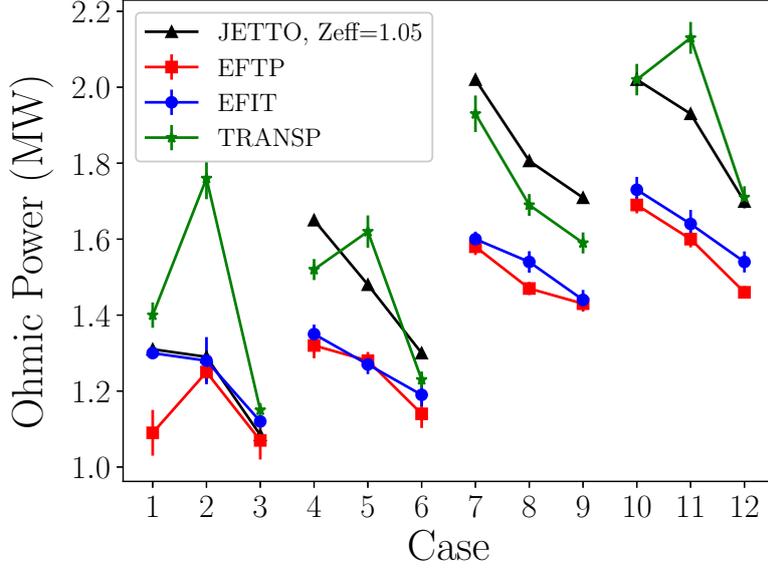


Figure 6.3: Values of the Ohmic power measured in EFIT, EFTP and TRANSP, with the corresponding JETTO interpretive simulations with  $Z_{\text{eff}} = 1.05$ . Here the cases refer to those in increasing order of table 6.1, such that cases 1, 2 and 3 refer to the H, D and T shots of the low-density set.

## 6.4 JETTO simulations

Each of the 15 cases (3 L-modes and 12 Ohmic discharges) were simulated in JETTO using 5 different models for the turbulence calculations, totalling 75 simulations. All cases were simulated once interpretively to find the experimental confinement time, and then once each predictively using 4 turbulence models, TGLF-SAT1, TGLF-SAT2, TGLF-SAT3 and QuaLiKiz-QLK.

### JETTO settings

For all simulations a single equilibrium construction obtained from EFIT was used, taken at 55.3 s for the L-modes and at the start of the time window presented in table 6.1 for the Ohmic pulses. The evolution of the current was stipulated to not change the equilibrium geometry for simplicity. 101 grid points were included in JETTO<sup>7</sup>, and runs were simulated for at least 1 second, which was generally sufficient for profiles to equilibrate<sup>8</sup>. A maximum time step of  $4.0 \times 10^{-3}$  s was used. Convergence tests were performed which demonstrated that results

<sup>7</sup>Note this is not the number of points at which the turbulent transport code is called, but the points at which all other models are evaluated.

<sup>8</sup>Simulations that had not were then continued past a time of 1 s until they had.

were unchanged by reducing this value.

The domain over which the turbulent models were simulated was up to  $\rho_J = 0.99$ , for JETTO normalised radial coordinate  $\rho_J$ . The temperature profiles used to initiate the simulations were taken from the input profiles and fixed at the  $\rho_J = 0.99$  flux surface. The typical radial grid size for the turbulence models was  $N_\rho = 30$ .

Beryllium  ${}^9_4\text{Be}$  impurities were included but not evolved predictively, through a radially constant  $Z_{\text{eff}} = 1.05$  for the Ohmic cases, and  $Z_{\text{eff}} = 1.1$  for the L-mode cases. The low levels of impurities observed experimentally justified these values. Neutrals were also not included, owing to the interpretive simulation of the density profiles.

The NBI present in the L-modes was simulated using the PENCIL code, with the beam parameters specified via beam boxes. The parameters for octants 4 and 8 of the different L-mode simulations are included in table 6.2. Note that octant 4 was not used in shot #99173 (T) over the time window considered. No pellet injection was included.

	Quantity (Octant 4, Octant 8)			
Pulse	Energy (keV)	Fractions	PINI's	Power (MW)
91450 (H)	61, 69	(0.31, 0.29, 0.39), (0.31, 0.33, 0.36)	(1, 2, 3, 4, 8), (7, 8)	3.00, 1.38
89723 (D)	83, 91	(0.5, 0.25, 0.25), (0.51, 0.27, 0.22)	(1, 6), (7)	2.09, 1.11
99173 (T)	-, 96.5	-, (0.55, 0.24, 0.21)	-, (6, 7)	-, 1.45

Table 6.2: Parameters for the NBI heating used in PENCIL for the L-mode simulations.

Neoclassical transport was simulated using NCLASS. A Gaussian transport profile was included in the core, approximating additional transport due to the sawtooth instability<sup>9</sup> [149]. Equivalent for particle and thermal transport, this is given by

$$D_G, \chi_G = H_G e^{-(\rho_J - c_G)^2 / w_G^2} \text{ cm}^2 \text{ s}^{-1} \quad (6.1)$$

where  $H_G = 10^4$ ,  $c_G = 0.1$  and  $w_G = 0.05$ . Lower limits were placed on the particle and thermal diffusion coefficients of  $3 \times 10^{-3} \text{ cm}^2 \text{ s}^{-1}$  for numerical stability.

<sup>9</sup>This is an MHD-type mode which involves a sharp drop in temperature and density in the tokamak core where  $q < 1$ , as particles and energy are transported radially during a ‘crash’. The profiles then increase again, approximately linearly with time, until another sawtooth crash occurs, with the cycle thus repeating. This quasi-periodic rising and crashing of the profiles gives the instability its name.

The effect of  $q$  profile pinning due to sawteeth was included via a continuous sawtooth model which sets  $q = 1$  as a lower bound for the  $q$  profile. Radiation was included for the L-modes with a radially-constant value of 0.5 MW, to account for the bolometry measurements made of the total radiated power. No radiation was included for the Ohmic cases. No fusion reactions were included.

### Turbulence code settings

Simulations involving TGLF used `KYGRID_MODEL = 4`. This corresponds to 12  $k_y \rho_i$  values in the ion scale, with 10 evenly spaced between  $k_y \rho_i = 0.1$  and  $k_y \rho_i = 1.0$ , and two additional modes at  $k_y \rho_i = 0.05$  and  $k_y \rho_i = 0.15$ . Note these grids are scaled to the main isotope. These additional modes are included to improve the resolution of the calculation of  $\gamma_{\max}$  and  $k_{\max}$ . 8  $k_y$  modes were simulated in the electron scale, logarithmically spaced between  $k_y \rho_i = 1.0$  and  $k_y \rho_e = 0.4$ , and the two fastest growing modes were simulated (`NMODES = 2`). The collision model `XNU_MODEL = 2` was used for SAT1 simulations, and `XNU_MODEL = 3` was used for SAT2 and SAT3, as these were the settings to which the saturation rules had been trained.

For the QuaLiKiz simulations version 2.8 was used, with all parameters set to their default values.

### Methodology

To quantify the overall predictive capability of these JETTO runs, we measure the energy confinement time of a converged simulation, equivalent to the confinement time introduced in equation 1.5,

$$\tau_E = W / \left( P_{\text{in}} - \frac{dW}{dt} \right) \quad (6.2)$$

where  $W$  is the stored energy of the plasma and  $P_{\text{in}}$  is the total input power, for which radiation is not included. The comparison between the predicted values of  $\tau_E$  from the different quasilinear models to that obtained from the interpretive simulation indicates the degree to which the JETTO simulation can predict the experimental confinement.

In particular, we are interested in the isotope scaling between the confinement times of shots in the same set. The isotope scaling parameter  $\alpha_A$  can again be used for this purpose, such that we fit data of the form  $\tau_E = C_A A^{\alpha_A}$ . Note, this is equivalent to the power law scaling value typically used in experiment, as in [61, 150]. However unlike with the gyrokinetic database of section 4.6, in the experiment several parameters are changed between shots, not just the

isotope mass, and so one does not necessarily expect the confinement time to scale with a single exponent between the 3 different isotope shots of a given set. Instead, two values of  $\alpha_A$  may be used, each describing the scaling between two isotopes, as in equation ???. Hence we introduce  $\alpha_{\text{HD}}$  to describe the scaling between H and D, and  $\alpha_{\text{DT}}$  for the scaling between D and T. A third may also be defined, quantifying the scaling between H and T,  $\alpha_{\text{HT}}$ , however the value of one of the three exponents is implied given knowledge of the other two. One may show that these scaling factors are related via

$$3^{\alpha_{\text{HT}}} = 2^{\alpha_{\text{HD}}} \left(\frac{3}{2}\right)^{\alpha_{\text{DT}}}. \quad (6.3)$$

Now applied to the confinement time, a measured  $\alpha_A$  value of  $-0.5$  corresponds to agreement with the naive gyroBohm scaling prediction, as discussed in section 2.6. A value close to zero implies no change in the confinement time with isotope, and a positive number is a manifestation of the isotope effect (anti-gyroBohm scaling).

#### 6.4.1 L-mode simulations

We first consider the results of the L-mode simulations. The plot of confinement time against pulse number is shown in figure 6.4a, with the measured isotope scaling metric values presented on the left-hand side of table 6.3. The predicted temperature profiles of SAT1, SAT2, SAT3 and QuaLiKiz along with the experimental profiles (with  $T_i = T_e$ ) are shown in figure 6.5.

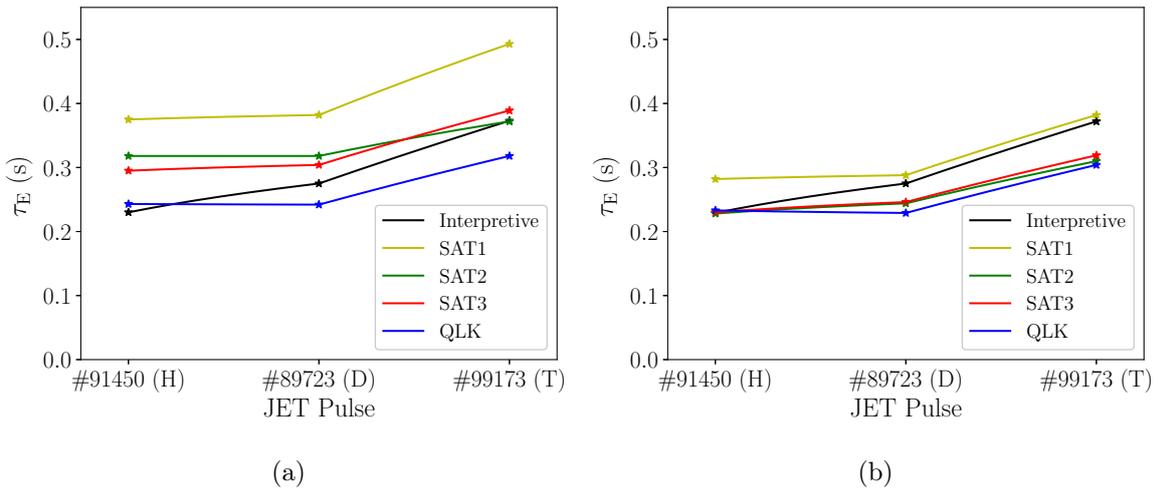


Figure 6.4: Confinement time  $\tau_E$  obtained for JETTO simulations using different quasilinear transport models for the set of L-mode pulses with (a) all effects included and (b) with no rotation.

	Full simulation			No rotation		
Model	$\alpha_{HD}$	$\alpha_{DT}$	$\alpha_{HT}$	$\alpha_{HD}$	$\alpha_{DT}$	$\alpha_{HT}$
Interpretive	0.26	0.75	0.44	0.26	0.75	0.44
SAT1	0.03	0.63	0.25	0.03	0.70	0.28
SAT2	0.00	0.39	0.14	0.10	0.59	0.28
SAT3	0.04	0.61	0.25	0.10	0.64	0.30
QLK	-0.01	0.67	0.24	-0.02	0.70	0.24

Table 6.3: Isotope scaling metric values between different shots for the L-mode set with all effects included on the left hand column, and no momentum transport on the right hand column.

Considering figure 6.4a, then for the interpretive simulations positive scaling is found across the set, exhibiting anti-gyroBohm scaling as expected. This result can be seen to be approximately recreated in all of the quasilinear models, namely the scaling is generally positive, indicating a recreation of the isotope effect in integrated modelling with all turbulence models. The agreement in the scaling is particularly good between the D and T simulations, however the scaling between H and D is too weak, with all of the predictive simulations measuring  $\alpha_{HD} \approx 0.0$ , whereas the interpretive simulation finds the scaling to be  $\alpha_{HD} = 0.26$ . The magnitudes of the confinement times are seen to be close for all models except SAT1, which overpredicts the confinement.

Turning to figure 6.5 then the temperature profiles of SAT2 and SAT3 are seen to be very similar, in general overpredicting all profiles other than the ion temperature in the tritium discharge. Given the similarity of the profiles from these models this may indicate that the TEM-branch of SAT3, the principal difference between the models, is not being activated significantly. QuaLiKiz is seen to perform well in the H discharge, with an underpredicted ion temperature in D and T.

Lack of isotope scaling between shots such as those predicted for H and D using the turbulence models can generally be attributed to profile stiffness<sup>10</sup> [151, 152]. However in the case of anti-gyroBohm scaling in predictive modelling, such as is observed between D and T, several effects

<sup>10</sup>This is the quality of turbulent transport for which there is minimal transport at low driving gradients until a critical gradient is reached, at which point the transport rises sharply with a small increase in gradient. The steepness of this dependency effectively ‘pins’ the profile close to this critical gradient, and thus similar levels of transport and confinement are produced.

have been put forward as candidates to explain this result. These include TEM-dominated turbulence [64, 133, 153],  $E \times B$  shear [154] and the electron-ion energy exchange  $P_{ei}$  [60, 147]. This is a term that appears in the energy transport equations and has a relation with the ion mass and temperature difference via  $P_{ei} \propto (T_e - T_i) / m_i$ . The observation that all models predict  $T_e > T_i$  for the T discharge (figures 6.5c, 6.5f) indicates the importance of this term for this case, motivating separate measurements of  $T_i$  for use in future work.

An advantage of integrated modelling is that these effects can be artificially turned off or scaled to investigate their relative influence on the confinement. For example, all the L-mode simulations were repeated with toroidal momentum transport (the effect of  $E \times B$  shear) turned off. The results of the confinement time with pulse number for this exercise is shown in figure 6.4b, with the  $\alpha_A$  values given on the right-hand side of table 6.3. While a reduction in the magnitudes of the confinement time is observed across the models, the isotope scaling remains relatively unchanged, indicating that here the  $E \times B$  shear is not responsible for the anti-gyroBohm scaling effect. Further investigation into the other potential mechanisms is required to better tie down the origin of this scaling.

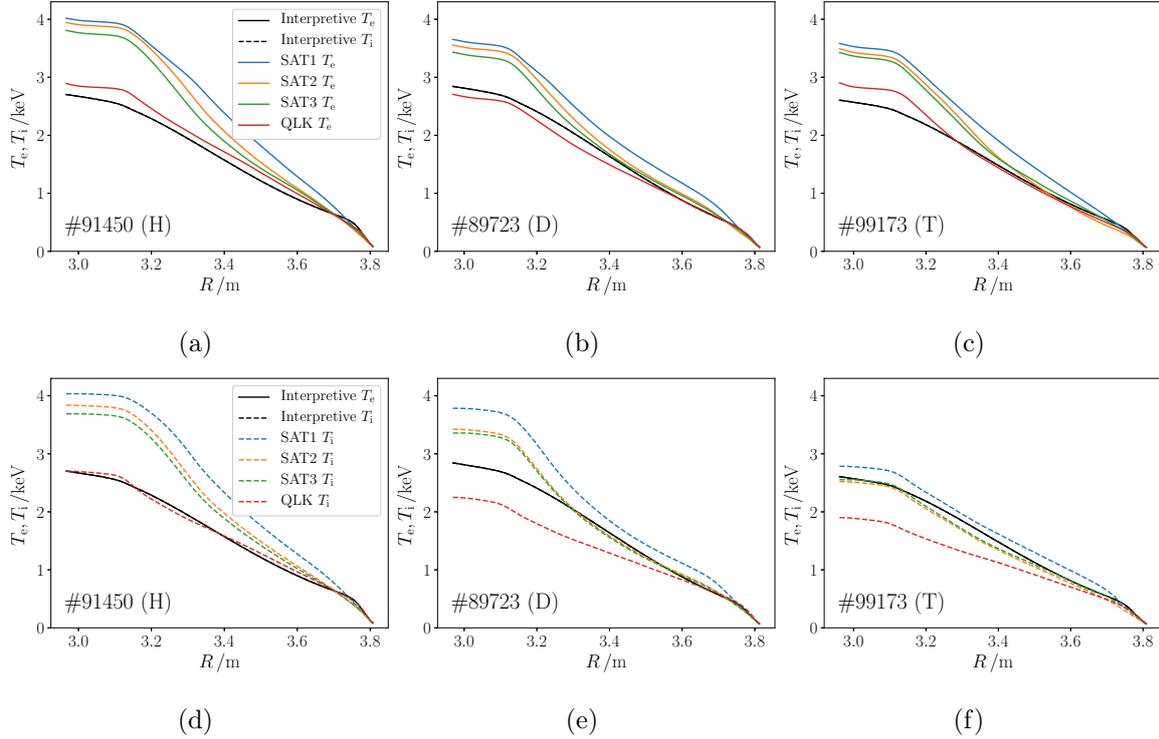


Figure 6.5: Predicted electron ((a)-(c)) and ion ((d)-(f)) temperature profiles against major radius for the 3 L-mode simulations using various transport models. Note  $T_i = T_e$  was assumed for the interpretive profiles.

## 6.4.2 Ohmic simulations

The confinement results of Ohmic sets 1 and 2 are shown in figure 6.6, with the corresponding  $\alpha_A$  values in table 6.4. For both cases the interpretive simulations again indicate anti-gyroBohm scaling. Compared to the L-modes however, here the agreement between the models and the interpretive simulations is seen to be less robust. In the low-density set the 3 TGLF models again describe the positive scaling trend, however are too strong between H and D and too weak between D and T. Better agreement is found for SAT2 and SAT3 in the mid-density set, however the magnitudes of  $\tau_E$  for all TGLF rules for both sets are generally overpredicted. For QuaLiKiz, anti-gyroBohm scaling is not produced, with a *negative* scaling between H and D, and H and T in both cases.

The results of the high-density sets 3 and 4 are shown in figure 6.7 and the isotope scaling  $\alpha_A$  shown in table 6.5, for which we note similar observations to sets 1 and 2.

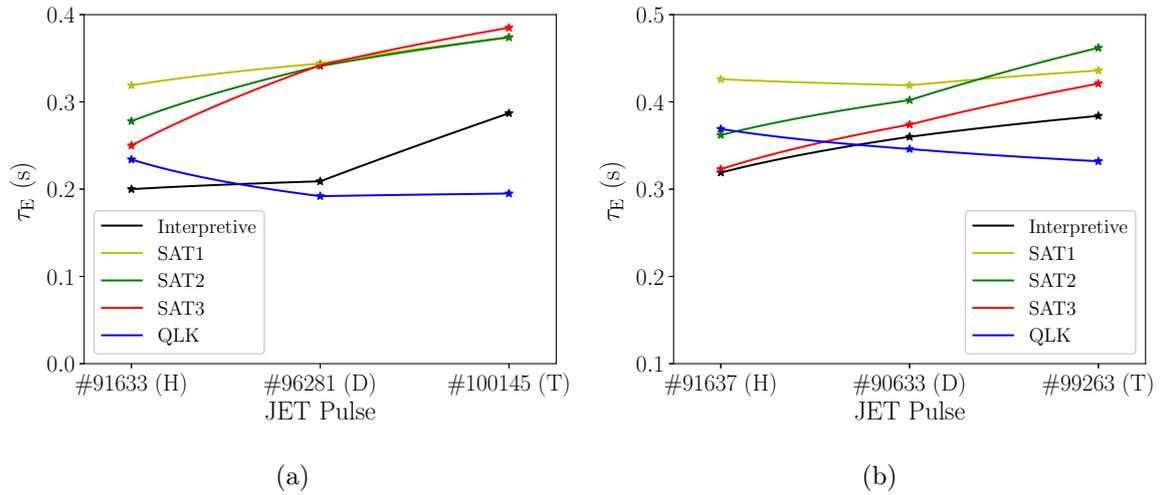


Figure 6.6: Confinement time  $\tau_E$  obtained for JETTO simulations using different quasilinear transport models, for a) the low-density set of Ohmic pulses and b) the mid-density set.

The ion and electron temperature profiles of the mid-density set,  $\bar{n}_e \sim 1.45 \times 10^{19} \text{ m}^{-3}$ , are shown in figure 6.8. The observed features here can be shown to be representative of the remaining Ohmic cases, namely that there is the largest overprediction in the temperature profiles for the H discharges, with the models coming closer into agreement with D and then T. This suggests the physical relevance of an isotope effect in these discharges that is increasingly well-captured with ion mass in the transport models, rather than these results being caused by the models' historical tuning to deuterium simulations. Note again the similarity in the performance of SAT2 and SAT3.

A recreation of the LOC-SOC plot of figure 6.2 for each turbulence model for the 12 Ohmic

	$\bar{n}_e \sim 0.87 \times 10^{19} \text{ m}^{-3}$ (Set 1)			$\bar{n}_e \sim 1.45 \times 10^{19} \text{ m}^{-3}$ (Set 2)		
Model	$\alpha_{\text{HD}}$	$\alpha_{\text{DT}}$	$\alpha_{\text{HT}}$	$\alpha_{\text{HD}}$	$\alpha_{\text{DT}}$	$\alpha_{\text{HT}}$
Interpretive	0.06	0.78	0.33	0.17	0.16	0.17
SAT1	0.11	0.21	0.14	-0.02	0.10	0.02
SAT2	0.29	0.23	0.27	0.15	0.34	0.22
SAT3	0.45	0.29	0.39	0.21	0.29	0.24
QLK	-0.29	0.04	-0.17	-0.09	-0.10	-0.10

Table 6.4: Isotope scaling metric values between different isotopes for the low- and mid-density Ohmic sets.

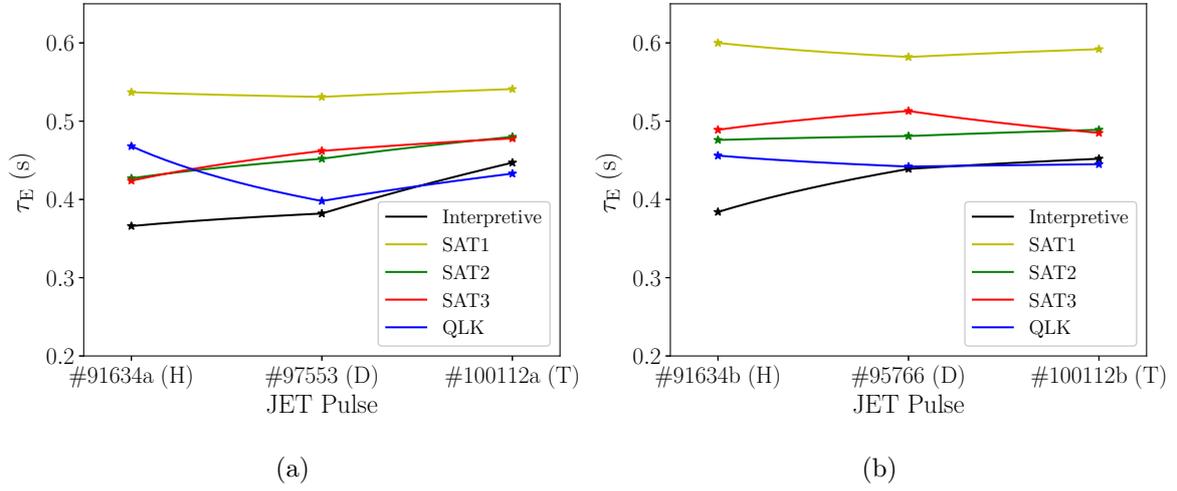


Figure 6.7: Confinement time  $\tau_E$  obtained for JETTO simulations using different quasilinear transport models, for a) the first and b) the second high-density Ohmic sets.

cases is shown in figure 6.9, for which some of the properties observed in the foregoing can be seen more distinctly. SAT1 generally overpredicts confinement, exhibits a small degree of isotope scaling and struggles to recreate the saturation of the SOC regime. SAT2 and SAT3 perform similarly, demonstrating anti-gyroBohm scaling across the majority of the scan and presenting a semblance of the higher-density saturation. Finally, QuaLiKiz performs the best in the scan for D, closely adhering to the experimental results across the density range. However, QuaLiKiz does not recreate the anti-gyroBohm scaling of the experiment, instead producing a gyroBohm-like scaling for H and T, indicating part of the description necessary for this anti-gyroBohm effect is missing in the physics of QuaLiKiz.

	$\bar{n}_e \sim 2.30 \times 10^{19} \text{ m}^{-3}$ (Set 3)			$\bar{n}_e \sim 2.80 \times 10^{19} \text{ m}^{-3}$ (Set 4)		
Model	$\alpha_{\text{HD}}$	$\alpha_{\text{DT}}$	$\alpha_{\text{HT}}$	$\alpha_{\text{HD}}$	$\alpha_{\text{DT}}$	$\alpha_{\text{HT}}$
Interpretive	0.06	0.39	0.18	0.19	0.07	0.15
SAT1	-0.02	0.05	0.01	-0.04	0.04	-0.01
SAT2	0.08	0.15	0.11	0.02	0.04	0.02
SAT3	0.12	0.08	0.11	0.07	-0.14	-0.01
QLK	-0.23	0.21	-0.07	-0.04	0.02	-0.02

Table 6.5: Isotope scaling metric values between different isotopes for the two high-density Ohmic sets.

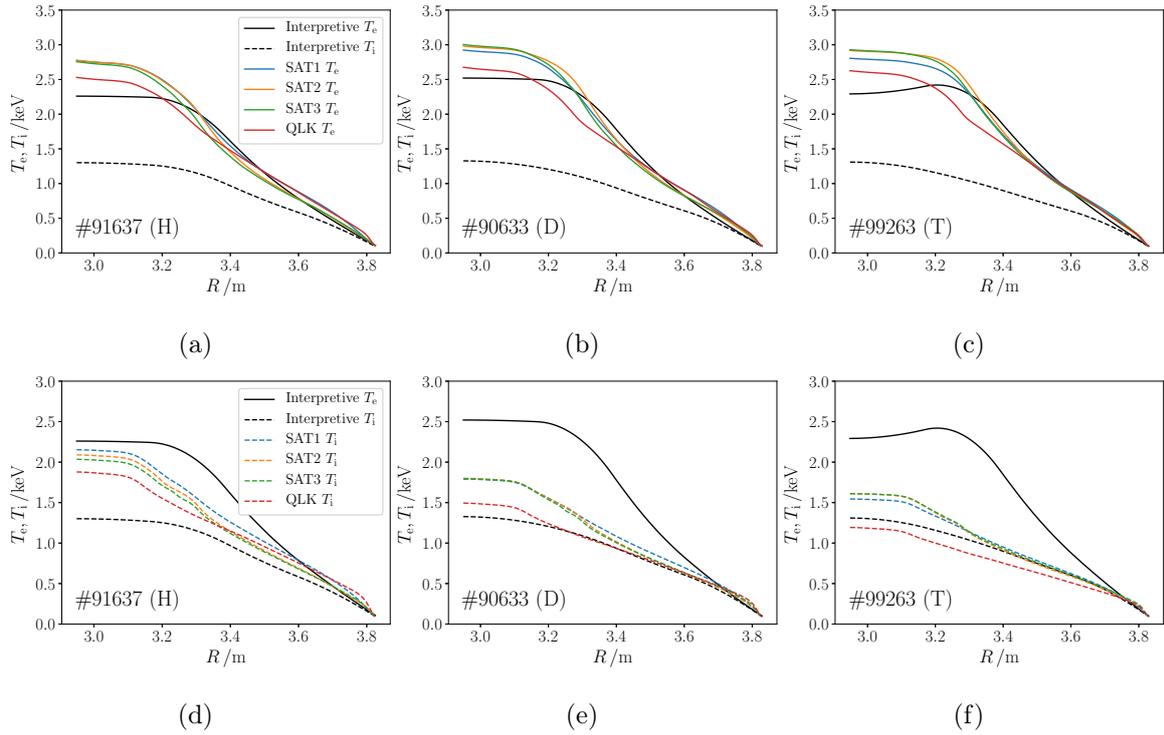


Figure 6.8: Electron ((a)-(c)) and ion ((d)-(f)) temperature profiles against major radius for predictive JETTO simulations of the mid-density Ohmic set, with the experimental profiles shown in black.

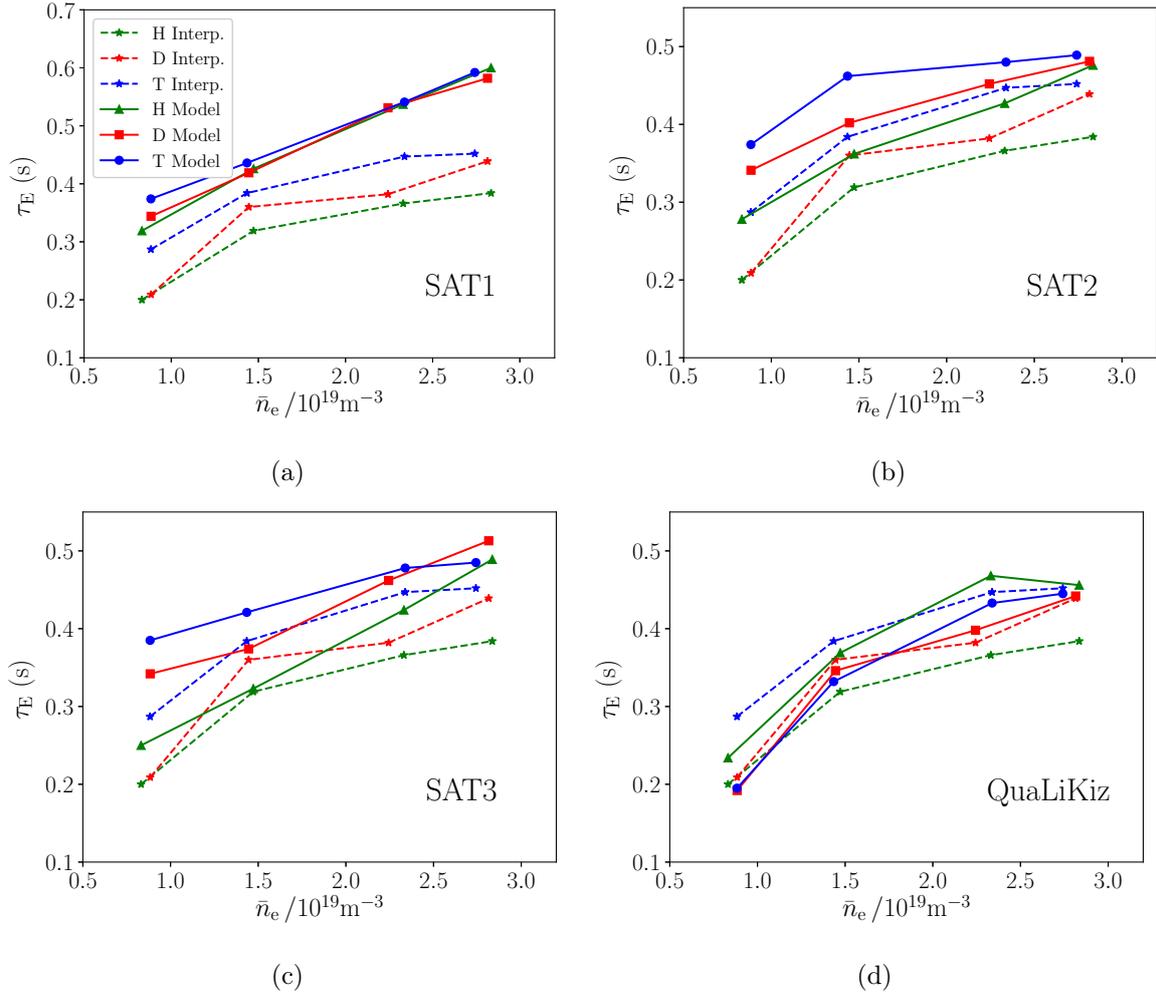


Figure 6.9: Confinement time  $\tau_E$  obtained for predictive JETTO simulations against line-averaged density  $\bar{n}_e$  for the four turbulence models relative to the interpretive results.

## 6.5 Summary and future works

The results of this chapter have shown that for the L-modes and Ohmic cases considered, the performance of SAT3 in integrated modelling produces similar results to those of SAT2. Both of these models perform well overall, able to generally recreate the anti-gyroBohm scaling of the confinement seen in experiment more accurately than SAT1 and QuaLiKiz. SAT2 and SAT3 were developed concurrently over the course of this PhD, with observations and results made for one rule typically informing the other. Indeed, the aspects of the training database of SAT3 that were not included ( $E \times B$  shear, electron scale, subdominant modes) were incorporated from SAT2, as discussed in section 5.5. This means that SAT3 can loosely be considered as model similar to SAT2, with the inclusion of a saturation rule branch for the modelling of TEM-dominated turbulence. The similarity of the results between SAT2 and SAT3 indicates that this branch is not having a significant influence on the evolution of the confinement for the simulations considered, potentially due to the model stiffness relaxing the profiles to the critical gradient regardless of the saturation branch.

To confirm and quantify the effect of the TEM-dominated branch of SAT3 in future works, several studies should be conducted. By measuring the values of  $\left| W_{e,k_y}^L / \sum_i W_{i,k_y}^L \right|_{k_y=k_{\max}}$  at each of the grid points using the dump files for the turbulence models in the integrated simulations, a radial profile of the relative strength of TEM-dominated turbulence can be produced. This can then be used to directly observe the extent of the region for which the SAT3 branch is acting. The direct effect of this branch can also be probed by re-running the simulations of this section with the TEM-dominated branch turned off.

The simulations should also be repeated with the density profile evolution modelled predictively, as here the evaluation of particle transport as calculated by the turbulence models has been neglected. The steep gradients in density at the edge of the plasma should trigger the TEM-dominated branch for SAT3 and thus produce disparate isotope scaling of the particle transport relative to SAT2, which may cause the calculation of the confinement time to be in greater contrast to the results of this chapter.

SAT1 was generally found to not perform as well as SAT2 and SAT3 in both the isotope scaling and magnitudes of the confinement times, exhibiting systematic overprediction in the latter. This may be a result of the free parameter associated with the saturation level of SAT1 being tuned to under-resolved nonlinear gyrokinetic simulations, and thus predicting a smaller level of transport.

The results of QuaLiKiz generally did not capture the isotope scaling of the confinement, in

the majority of cases finding a negative (gyroBohm) isotope scaling closer to that suggested by dimensional arguments rather than the positive scaling observed experimentally. The measured confinement times were generally of a similar value to the interpretive cases, performing particularly well for the D simulations.

# Chapter 7

## Summary and discussion

The existence of microinstabilities in tokamak plasmas gives rise to the turbulent transport of energy and particles, forming the dominant loss mechanism in a tokamak and limiting confinement. The use of quasilinear turbulence models to calculate this transport in integrated modelling simulations allows for physics-based predictions of the confinement properties of experimental plasma discharges, however the inherent simplifications of these models and their tuning to nonlinear gyrokinetic results can render them unreliable in less-explored parameter regimes, compromising the ability to perform effective integrated modelling studies. In preparation for ITER and the future shift from experimental discharges of pure D plasmas to a mixture of D-T, the question of quasilinear turbulent transport must be considered in the context of the anti-gyroBohm scaling of confinement, which this thesis has sought to address. This has been done in three main parts, forming a cycle of quasilinear model development as presented in section 2.5.1:

1. Comparison between current quasilinear models and nonlinear gyrokinetic simulations, including the diagnosis of the origin of the anti-gyroBohm scaling of transport (Chapter 4).
2. The development of a new saturation model, verified to be able to recreate the isotope scaling of fluxes seen in nonlinear gyrokinetics (Chapter 5).
3. The validation of the new saturation model in integrated modelling simulations against experimental data (Chapter 6).

This work forms part of a body of contemporary research investigating the isotope effect more generally, which is itself embedded in the field of tokamak confinement, working towards a greater understanding of the viable modes of operation of future tokamaks. The new results

obtained in the three parts of this work will now be discussed, including their limitations and possible directions of future research.

### **Comparison between quasilinear models and nonlinear gyrokinetics**

The analysis of Chapter 4 confirmed that existing quasilinear models generally struggle to capture the isotope scaling of fluxes seen in nonlinear gyrokinetic simulations, in part due to the historical lack of focus on isotopes other than deuterium in the development of their saturation rules and an incomplete understanding of the isotope effect. Within the quasilinear decomposition, the deviation from the gyroBohm scaling of fluxes was demonstrated to originate primarily in the magnitude of the saturated electrostatic potentials, particularly in the region of low  $k_y$  around the peak of the potential spectrum. This is due to its greater relative contribution to the overall transport and the variation in isotope scaling seen in this region, observed to be a common effect in both ITG- and TEM-dominated regimes. The feature that differentiates the two turbulence regimes is the shift from gyroBohm scaling common to all values of  $k_y$ , with those present in TEM-dominated cases being more negative than those of ITG turbulence, such that anti-gyroBohm scaling is produced in TEM cases. Current quasilinear models fail to replicate this behaviour due to missing the correct offset and/or the low  $k_y$  isotope variation in their saturation rules.

This observed deviation in scaling originates physically from the non-adiabatic part of the electron distribution, however it is primarily manifested in the nonlinear saturation, rather than the phase between the potential and plasma fluctuations. This answers the question posed in [64] regarding whether the reduced electron equations used in quasilinear fluid models could in principle be used to accurately describe the isotope scaling of fluxes, given their non-precise treatment of the non-adiabatic dynamics. The results of this study suggest they can if the effect is captured in the saturation rule, rather than being a direct remit of the dynamical equations of the electrons.

The analysis of this chapter was facilitated by computing the scaling exponent of various quantities with mass  $\alpha_A$ . This metric can be used to present the scaling information of three isotope cases in H, D and T as a single intuitive number, simplifying the ability to interpret scaling behaviour as compared to ratios. Initially used for the 0D turbulent flux, the use of  $\alpha_A$  was then generalised to allow description of the isotope scaling of one dimensional quantities<sup>1</sup>, including the flux components and saturated potential magnitudes, from which the causes of deviation from gyroBohm scaling in both gyrokinetics and quasilinear models could be clearly

---

<sup>1</sup>In principle this can be extended further, with  $\alpha_A$  analysis of 2D quantities able to produce an ‘ $\alpha_A$ -surface’.

presented. A limitation of this metric can arise in the case that the scaling is not sufficiently uniform across the scan, such as is found in the confinement times of the integrated modelling discharges of Chapter 6. In this case, the scaling can be exactly described using two numbers (e.g.  $\alpha_{\text{HD}}$  and  $\alpha_{\text{DT}}$ ) as opposed to one, in line with previous scaling metrics. Given that an  $\alpha_A$  value can be calculated for all previous metrics used in the literature, it represents a more flexible generalisation of the description of isotope scaling and its use is encouraged in future works.

During a convergence study for the nonlinear gyrokinetic database it was observed that the binormal wavenumber resolutions typical of ion-scale turbulence studies,  $N_y \sim 16$ , can under-resolve the fluxes by  $\sim 25\%$ , motivating the use of  $N_y = 40$ . This observation should encourage a reconsideration of the convergence criteria of ion-scale nonlinear gyrokinetic simulations.

### **New saturation rule development**

In Chapter 5 the new quasilinear saturation rule SAT3 was derived, in part informed by the observations of the previous chapter. The form of the 2D potential spectrum in  $k_x$  and  $k_y$  was first considered, generalising the previous assumption of a squared Lorentzian function in the radial wavenumber through a self-consistent description in terms of the first three  $k_x$ -moments of the spectrum. The 2D model adhered to the data extremely well across a range of decades in radial wavenumber.

A robust novel relationship between the saturated potential for each  $k_y$  integrated over  $k_x$  and the radial spectral widths was discovered in the nonlinear gyrokinetic simulation results, which appears to capture a conserved quantity in electrostatic turbulence. This forms the basis of the spectral shape of SAT3, however the understanding of the physical mechanism responsible for this relation, as well as that for the cascade scales predicted for the 2D spectrum, is consigned to future work.

The database demonstrated different saturation levels for ITG- and TEM-dominated turbulence from the scans in isotope<sup>2</sup>, which motivated a saturation model for each mode type with a transition function between them. Generalisations of the model outside of the dataset were implemented, namely the effect of  $E \times B$  shear, multi-ion plasma operation and electron-scale saturation by connecting with previous work. This brought the model closely in line with the state-of-the-art SAT2, however now with the ability to capture the saturation level of TEM-dominated turbulence, an improved spectral shape and a normalisation based on better

---

<sup>2</sup>To investigate the generality of this difference in other parameters requires additional TEM-dominated simulations.

converged nonlinear gyrokinetic fluxes. SAT3 therefore represents an incremental generalisation of a previously established model, and is a tool that is of general use to the integrated modelling community.

SAT3 was verified to better capture the isotope scaling of the cases in the database, particularly in the cases of TEM-dominated turbulence, while performing at least as well as existing quasilinear models in other parameters. Having been constructed from a database of electrostatic core plasmas, dominated by either ITG or TEM turbulence, SAT3's applicability is naturally most valid in these regions of parameter space. Caution should therefore be exercised in the use of SAT3 away from these areas, such as plasmas for which electromagnetic effects are expected to be significant or in the presence of dominant modes other than the ITG or TEM.

Future generalisation of SAT3 into these regimes will require comparison to additional nonlinear gyrokinetic simulations. The methodology used in this work to derive SAT3 represents an algorithm which can effectively be employed in such studies, based on the systematic comparison of the constituent parts of quasilinear models to nonlinear gyrokinetic data, working via a 'bottom-up' approach. This can be applied to both the improvement of current saturation rules and to the generation of new ones.

### **Experimental validation of the new saturation model**

In Chapter 6 SAT3 was incorporated into JETTO for validation against experimental data in H, D and T from JET, examining three L-mode discharges and 12 Ohmic discharges. The results of SAT3 were also compared with other contemporary quasilinear models, namely SAT1, SAT2 and QuaLiKiz. The work presented in this chapter represents the initial stages of a more thorough study into the validation of SAT3 and of the isotope effect in integrated modelling, to be continued in future work.

SAT3 was seen to produce similar results to SAT2, both of which performed well relative to the experimental profiles. They both consistently captured the anti-gyroBohm scaling of the confinement time, generally with slight overpredictions in the magnitudes. The similarity between the two models indicates that the influence of the TEM-dominated saturation branch of SAT3 was not particularly strong. Future studies will verify this through explicitly measuring over what portion of the minor radius this TEM-dominated branch is triggered, as well as conduct these simulations using predictive density, so as to allow the equilibria to relax to a 'consistent' profile. It is here that the TEM-dominated branch of SAT3 could play a more prominent role in moving the SAT3 result to a more distinct equilibrium than that of SAT2.

As well as ascertaining the influence of SAT3’s TEM-dominated branch, further studies should be conducted into the relative effects of other gyroBohm-breaking effects, to determine the origin of the appearance of anti-gyroBohm scaling in integrated modelling simulations. Initial stages of this study were conducted, for which it was found that neglecting toroidal momentum transport in the L-mode cases minimally changed the isotope scaling of the models, indicating its lack of influence in breaking the gyroBohm scaling for the relatively low-rotation cases studied here. By controlling the presence of other effects, such as turning off the TEM-dominated branch of SAT3, turning off the multiscale saturation in the models, as well as altering the electron-ion energy exchange term  $P_{ei}$  in the energy transport equations, the importance of each of these for the isotope effect in integrated modelling can be measured.

### **Future quasilinear model development**

Having addressed the behaviour of quasilinear models with isotope mass, there are several other timely considerations for future works, which are presented here. In these cases a cycle of quasilinear development is envisioned, including comparison to nonlinear gyrokinetic results, subsequent improvement of the quasilinear models and validation with experimental data.

One such area is that of mixed plasmas, due to its importance for the future operations of D-T. While current quasilinear models are able to in principle operate with multiple ion species, a dedicated comparison with the results of nonlinear gyrokinetics has not yet taken place. Additional complications with respect to a pure ion plasma are anticipated, including the effect of collisions between the two ion species, as well as the interplay of two turbulent modes of the same type and of similar growth rate, such as the simultaneous presence of a ‘D’ ITG and a ‘T’ ITG for a D-T mix. Whether such a plasma acts similarly to a pure plasma with an ‘effective mass’ related to the number densities of the two,  $m_{\text{eff}} \sim (m_1 n_1 + m_2 n_2) / (n_1 + n_2)$ , or whether distinct effects are incurred should be investigated, as well as the ability of current quasilinear models to replicate them. The recent JET isotope experiments contained mixture discharges of D-T as well as H-T of various concentrations, providing experimental data with which to validate models in these studies.

Another consideration is that of helium plasmas. These will form part of ITER’s Pre-Fusion Power Operation phase [155], with JET undertaking a preparatory helium campaign in late 2022. Helium plasmas have a non-hydrogenic charge of  $Z = 2$ , and thus presents a novel parameter space with which to investigate the performance of quasilinear models through charge and mass effects [67]. The JET campaign will produce experimental data for model validation purposes.

Both of these considerations are now more feasible as a result of the increase in understanding of the isotope dependence of turbulence from this work, as they vary both isotope mass and another parameter, namely number density for mixed plasmas and charge for helium.

A third and more challenging consideration is the development of an electromagnetic saturation rule using the model development methodology presented in this thesis, able to calculate the levels of transport resulting from kinetic ballooning mode and microtearing mode turbulence. This is an essential requirement for promoting confidence in integrated modelling simulations of spherical tokamaks, such as NSTX, MAST-U and STEP<sup>3</sup>. Basic electromagnetic quasilinear models have already been developed [156, 157], however not to the same level of sophistication as those describing electrostatic turbulence, due to the task of obtaining converged nonlinear simulations being notoriously challenging. Recently however, inroads have been made on the understanding of the conditions required, such that a preliminary database of microtearing and kinetic ballooning turbulence could be assembled, and a quasilinear model developed as per the algorithm used in this work. This would start with analysis of the underlying functional forms of the spectra, for example testing the applicability of equation 5.11 to the electromagnetic case. Previous observations that the nonlinear spectra of  $\delta\phi$  and  $\delta A_{\parallel}$  have a similar spectral shape [156] are an encouraging starting point for these studies.

---

<sup>3</sup>This is the UK's flagship spherical tokamak energy program, which aims to generate net electricity from fusion. A 'concept design' is anticipated by 2024, with construction aiming for completion around 2040.

# Appendix A

## Background concepts

### A.1 Velocity moments of the kinetic equation

Starting from the kinetic equation,

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \nabla f_s + \frac{q_s}{m_s} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_s}{\partial \mathbf{v}} = C_s \quad (\text{A.1})$$

one multiplies by a velocity dependent function  $g(\mathbf{v})$ , and integrates over all velocities. Here we shall use that the integral of the distribution function is defined as the number density of the species,

$$n_s(\mathbf{r}, t) = \int f_s(\mathbf{r}, \mathbf{v}, t) d^3\mathbf{v}. \quad (\text{A.2})$$

Because the velocity integrals are taken at constant  $\mathbf{r}$  and  $t$ , the integration commutes with  $\nabla$  and  $\partial/\partial t$ . Additionally, as  $g$  is only a function of  $\mathbf{v}$ , we have  $\nabla g = \mathbf{0}$  and  $\partial g/\partial t = 0$ .

We define the average of  $g$  over velocity space weighted by the distribution function via

$$\{g\}(\mathbf{r}, t) = \frac{\int f_s g d^3\mathbf{v}}{\int f_s d^3\mathbf{v}} = \frac{1}{n_s(\mathbf{r}, t)} \int f_s(t, \mathbf{r}, \mathbf{v}) g(\mathbf{v}) d^3\mathbf{v}. \quad (\text{A.3})$$

#### A.1.1 General velocity moment

For a given function of velocity  $g$ , the moment of the kinetic equation is

$$\int \frac{\partial f_s}{\partial t} g d^3\mathbf{v} + \int (\mathbf{v} \cdot \nabla f_s) g d^3\mathbf{v} + \int \left( \frac{q_s}{m_s} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_s}{\partial \mathbf{v}} \right) g d^3\mathbf{v} = \int C_s g d^3\mathbf{v}. \quad (\text{A.4})$$

The first term evaluates to

$$\begin{aligned} \int \frac{\partial f_s}{\partial t} g(\mathbf{v}) d^3\mathbf{v} &= \int \frac{\partial (f_s g)}{\partial t} d^3\mathbf{v} \\ &= \frac{\partial}{\partial t} \left( \int f_s g d^3\mathbf{v} \right) \\ &= \frac{\partial}{\partial t} (n_s \{g\}). \end{aligned} \quad (\text{A.5})$$

The second term is

$$\begin{aligned}
\int (\mathbf{v} \cdot \nabla f_s) g d^3\mathbf{v} &= \int \nabla \cdot (f_s g \mathbf{v}) d^3\mathbf{v} \\
&= \nabla \cdot \left( \int f_s g \mathbf{v} d^3\mathbf{v} \right) \\
&= \nabla \cdot (n_s \{ \mathbf{v} g \}).
\end{aligned} \tag{A.6}$$

The third and fourth terms require integration by parts. For the third term we consider the object

$$\frac{\partial}{\partial \mathbf{v}} \left( \frac{q_s}{m_s} f_s g \mathbf{E} \right) = \frac{q_s}{m_s} \mathbf{E} \cdot \frac{\partial f_s}{\partial \mathbf{v}} g + \frac{q_s}{m_s} \mathbf{E} \cdot \frac{\partial g}{\partial \mathbf{v}} f_s. \tag{A.7}$$

Integrating A.7 over all velocities, the left-hand side goes to zero, and we see the first term on the right-hand side corresponds to the third term in the kinetic equation. Therefore we have

$$\begin{aligned}
\int \frac{q_s}{m_s} \mathbf{E} \cdot \frac{\partial f_s}{\partial \mathbf{v}} g d^3\mathbf{v} &= - \int \frac{q_s}{m_s} \mathbf{E} \cdot \frac{\partial g}{\partial \mathbf{v}} f_s d^3\mathbf{v} \\
&= - \frac{q_s}{m_s} n_s \mathbf{E} \cdot \left\{ \frac{\partial g}{\partial \mathbf{v}} \right\}.
\end{aligned} \tag{A.8}$$

Similarly for the fourth term in equation A.4, using the result  $(\partial/\partial \mathbf{v}) \cdot (\mathbf{v} \times \mathbf{B}) = 0$ , we consider the object

$$\frac{\partial}{\partial \mathbf{v}} \left( \frac{q_s}{m_s} f_s g (\mathbf{v} \times \mathbf{B}) \right) = \frac{q_s}{m_s} g (\mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_s}{\partial \mathbf{v}} + \frac{q_s}{m_s} f_s (\mathbf{v} \times \mathbf{B}) \cdot \frac{\partial g}{\partial \mathbf{v}} \tag{A.9}$$

which upon integrating over all velocities, yields

$$\begin{aligned}
\int \frac{q_s}{m_s} g (\mathbf{v} \times \mathbf{B}) \cdot \frac{\partial f_s}{\partial \mathbf{v}} d^3\mathbf{v} &= - \int \frac{q_s}{m_s} f_s (\mathbf{v} \times \mathbf{B}) \cdot \frac{\partial g}{\partial \mathbf{v}} d^3\mathbf{v} \\
&= - \frac{q_s}{m_s} n_s \left\{ (\mathbf{v} \times \mathbf{B}) \cdot \frac{\partial g}{\partial \mathbf{v}} \right\}.
\end{aligned} \tag{A.10}$$

Bringing these results together, then equation A.4 becomes

$$\frac{\partial}{\partial t} (n_s \{g\}) + \nabla \cdot (n_s \{ \mathbf{v} g \}) - \frac{q_s}{m_s} n_s \mathbf{E} \cdot \left\{ \frac{\partial g}{\partial \mathbf{v}} \right\} - \frac{q_s}{m_s} n_s \left\{ (\mathbf{v} \times \mathbf{B}) \cdot \frac{\partial g}{\partial \mathbf{v}} \right\} = \int g C_s d^3\mathbf{v}. \tag{A.11}$$

## A.1.2 The evolution of density and momentum

### Density evolution

The first fluid equation is obtained using the zeroth moment,  $g = 1$ . From equation A.11 we obtain

$$\frac{\partial n_s}{\partial t} + \nabla \cdot (n_s \mathbf{u}_s) = 0 \tag{A.12}$$

where the integral of the collision term over all velocities is zero [158], and we have used the definition of the average fluid velocity, equation 1.20. Note in Cartesian components this is written

$$\frac{\partial n_s}{\partial t} + \sum_j \frac{\partial}{\partial x_j} (n_s u_j) = 0. \tag{A.13}$$

## Momentum conservation

We now consider the first moment. To bypass the use of dyadics in the second term, we shall first consider a Cartesian component of the momentum,  $g = m_s v_i$ , for  $i = x, y, z$ . Equation A.11 becomes

$$\frac{\partial}{\partial t} (n_s m_s u_{s,i}) + \nabla \cdot (n_s m_s \{\mathbf{v}v_i\}) - q_s n_s E_i - q_s n_s (\mathbf{u}_s \times \mathbf{B})_i = \int m_s v_i C_s d^3\mathbf{v}. \quad (\text{A.14})$$

Introducing the peculiar velocity  $\mathbf{c}_s = \mathbf{v} - \mathbf{u}_s$  and expanding the second term, we have

$$\begin{aligned} \{\mathbf{v}v_i\} &= \{(\mathbf{u}_s + \mathbf{c}_s)(u_{s,i} + c_{s,i})\} \\ &= \mathbf{u}_s u_{s,i} + \{\mathbf{c}_s c_{s,i}\} \\ &= \sum_j u_{s,i} u_{s,j} \hat{\mathbf{x}}_j + \sum_j \{c_{s,i} c_{s,j}\} \hat{\mathbf{x}}_j \end{aligned} \quad (\text{A.15})$$

where  $\hat{\mathbf{x}}_j$  are the basis vectors of the Cartesian system. The second term in equation A.14 is therefore

$$\nabla \cdot (n_s m_s \{\mathbf{v}v_i\}) = \sum_j \frac{\partial}{\partial x_j} (n_s m_s u_{s,i} u_{s,j}) + \sum_j \frac{\partial}{\partial x_j} (n_s m_s \{c_{s,i} c_{s,j}\}). \quad (\text{A.16})$$

Multiplying equation A.14 by the corresponding unit vector  $\hat{\mathbf{x}}_i$  and summing across the three dimensions, one has

$$\begin{aligned} \frac{\partial}{\partial t} (n_s m_s \mathbf{u}_s) + \sum_i \hat{\mathbf{x}}_i \left[ \sum_j \frac{\partial}{\partial x_j} (n_s m_s u_{s,i} u_{s,j}) \right] + \sum_i \hat{\mathbf{x}}_i \left[ \sum_j \frac{\partial}{\partial x_j} (n_s m_s \{c_{s,i} c_{s,j}\}) \right] \\ - q_s n_s (\mathbf{E} + (\mathbf{u}_s \times \mathbf{B})) = \int m_s \mathbf{v} C_s d^3\mathbf{v} \end{aligned} \quad (\text{A.17})$$

and so we have split the second term into a mean flow component and a peculiar velocity component.

The first two terms of equation A.17 can be combined using particle conservation. For the  $i^{\text{th}}$  Cartesian component, then using equation A.13 the sum of these two terms is

$$\begin{aligned} \frac{\partial}{\partial t} (n_s m_s u_{s,i}) + \sum_j \frac{\partial}{\partial x_j} (n_s m_s u_{s,i} u_{s,j}) &= \frac{\partial n_s}{\partial t} m_s u_{s,i} + \frac{\partial u_{s,i}}{\partial t} m_s n_s + \sum_j \frac{\partial}{\partial x_j} (n_s m_s u_{s,i} u_{s,j}) \\ &= \frac{\partial u_{s,i}}{\partial t} m_s n_s + m_s \sum_j \frac{\partial}{\partial x_j} (n_s u_{s,i} u_{s,j}) - u_{s,i} \frac{\partial}{\partial x_j} (n_s u_{s,j}) \\ &= \frac{\partial u_{s,i}}{\partial t} m_s n_s + m_s \sum_j n_s u_{s,j} \frac{\partial u_{s,i}}{\partial x_j} \\ &= \frac{\partial u_{s,i}}{\partial t} m_s n_s + m_s n_s (\mathbf{u}_s \cdot \nabla) u_{s,i} \end{aligned} \quad (\text{A.18})$$

and hence we have

$$\frac{\partial}{\partial t} (n_s m_s \mathbf{u}_s) + \sum_i \hat{\mathbf{x}}_i \left[ \sum_j \frac{\partial}{\partial x_j} (n_s m_s u_{s,i} u_{s,j}) \right] = m_s n_s \left[ \frac{\partial \mathbf{u}_s}{\partial t} + (\mathbf{u}_s \cdot \nabla) \mathbf{u}_s \right] \quad (\text{A.19})$$

In dealing with the third term of equation A.17, we define the *pressure tensor*

$$\mathbf{P}_{ij} = n_s m_s \begin{pmatrix} \{c_x^2\} & \{c_x c_y\} & \{c_x c_z\} \\ \{c_x c_y\} & \{c_y^2\} & \{c_y c_z\} \\ \{c_x c_z\} & \{c_y c_z\} & \{c_z^2\} \end{pmatrix} \quad (\text{A.20})$$

allowing the term to be written

$$\sum_i \hat{\mathbf{x}}_i \left[ \sum_j \frac{\partial}{\partial x_j} (n_s m_s \{c_{s,j} c_{s,i}\}) \right] = \sum_i \left[ \sum_j \frac{\partial P_{ij}}{\partial x_j} \right] \hat{\mathbf{x}}_i. \quad (\text{A.21})$$

This expression is the Cartesian component form of the divergence of a tensor field, with symbolic notation

$$\nabla \cdot \underline{\mathbf{P}} = \sum_i \left[ \sum_j \frac{\partial P_{ij}}{\partial x_j} \right] \hat{\mathbf{x}}_i. \quad (\text{A.22})$$

The pressure tensor is symmetric.

Noting that half of the sum of the diagonal is the peculiar kinetic energy density, we may *define* a scalar pressure,  $P_s$ , in accordance with the standard idea of pressure,

$$\begin{aligned} \frac{3}{2} P_s &= \frac{1}{2} n_s m_s (\{c_x^2\} + \{c_y^2\} + \{c_z^2\}) \\ &= \frac{1}{2} n_s m_s \{|\mathbf{c}|^2\} \\ &= \frac{1}{2} \text{Tr}(\mathbf{P}_{ij}) \end{aligned} \quad (\text{A.23})$$

One may also therefore define the species temperature as the ratio of the scalar pressure to the number density,

$$T_s = \frac{P_s}{n_s} = \frac{1}{3} m_s \{|\mathbf{c}|^2\}. \quad (\text{A.24})$$

We then split the pressure tensor into a scalar part and a viscosity part,  $\mathbf{P}_{ij} = P_s \delta_{ij} + \pi_{ij}$ .

Applying this separation to our term in the momentum equation, we have

$$\begin{aligned} \sum_i \sum_j \frac{\partial P_{ij}}{\partial x_j} \hat{\mathbf{x}}_i &= \sum_i \sum_j \frac{\partial P_s}{\partial x_j} \delta_{ij} \hat{\mathbf{e}}_i + \sum_i \sum_j \frac{\partial \pi_{ij}}{\partial x_j} \hat{\mathbf{e}}_i \\ &= \sum_i \frac{\partial P_s}{\partial x_i} \hat{\mathbf{x}}_i + \sum_i \sum_j \frac{\partial \pi_{ij}}{\partial x_j} \hat{\mathbf{x}}_i \\ &= \nabla P_s + \nabla \cdot \underline{\underline{\pi}} \end{aligned} \quad (\text{A.25})$$

and so we have separated the scalar pressure and the anisotropic pressure tensor. Bringing all these results together, the momentum equation is

$$m_s n_s \left[ \frac{\partial \mathbf{u}_s}{\partial t} + (\mathbf{u}_s \cdot \nabla) \mathbf{u}_s \right] + \nabla P_s + \nabla \cdot \underline{\underline{\pi}} - q_s n_s (\mathbf{E} + (\mathbf{u}_s \times \mathbf{B})) = \int m_s \mathbf{v} C_s d^3 \mathbf{v}. \quad (\text{A.26})$$

## A.2 Quasineutrality

Given the symmetry of the problem and the form of the potential of a point charge in vacuum, we consider the charge to be a function of radial distance only,  $\phi = \phi(r)$ . Gauss's law in the electrostatic case is  $\nabla^2\phi = -\rho/\epsilon_0$ , hence in spherical polar coordinates, one has

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\phi}{dr} \right) + \frac{1}{\epsilon_0} \rho = 0. \quad (\text{A.27})$$

Assuming a hydrogenic plasma with adiabatic electrons and stationary ions, the charge density becomes  $\rho = e(n_{i,0} - n_{e,0} - e\phi n_{e,0}/T_e) = -e^2\phi n_{e,0}/T_e$ , owing to the bulk neutrality of the plasma,  $n_{i,0} = n_{e,0}$ . Gauss's law becomes

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\phi}{dr} \right) - \left( \frac{e^2 n_{e,0}}{\epsilon_0 T_e} \right) \phi(r) = 0. \quad (\text{A.28})$$

For which the solution is of the form

$$\phi(r) = \frac{1}{r} \left( A e^{r/\lambda_D} + B e^{-r/\lambda_D} \right) \quad (\text{A.29})$$

where  $\lambda_D = \sqrt{\epsilon_0 T_e / e^2 n_{e,0}}$  is the Debye length. Imposing the boundary conditions of  $\lim_{r \rightarrow \infty} \phi \rightarrow 0$  ( $\implies A = 0$ ), and  $\lim_{r \rightarrow 0} \phi \rightarrow e / (4\pi\epsilon_0 r)$  ( $\implies B = e/4\pi\epsilon_0$ ), the point charge potential in vacuum, we get

$$\phi(r) = \frac{e}{4\pi\epsilon_0} \frac{e^{-r/\lambda_D}}{r}. \quad (\text{A.30})$$

## A.3 Phase in Fourier modes

We first show that phase difference is associated with the imaginary part of a Fourier amplitude. Consider the oscillation of a single mode of a real function. The condition for a function to be real is  $\hat{f}_{-n} = \hat{f}_n^*$ , and so for a single mode we have

$$f(t) = \hat{f}_n e^{i\omega t} + \hat{f}_n^* e^{-i\omega t}. \quad (\text{A.31})$$

Now using a polar representation for the generally complex Fourier amplitudes,  $\hat{f}_n = |\hat{f}_n| e^{i\theta_n}$ , we get

$$\begin{aligned} f(t) &= |\hat{f}_n| \left( e^{i(\omega t + \theta_n)} + e^{-i(\omega t + \theta_n)} \right) \\ &= 2 |\hat{f}_n| \cos(\omega t + \theta_n) \end{aligned} \quad (\text{A.32})$$

and thus we see that the phase of the oscillations  $\theta_n$  is associated with the imaginary part of the Fourier amplitudes.

Now consider two quantities  $f$  and  $g$  oscillating at the same frequency  $\omega_n$  but with different phases, such that their Fourier amplitudes are  $\hat{f}_n = |\hat{f}_n| e^{i\theta_n}$ ,  $\hat{g}_n = |\hat{g}_n| e^{i\phi_n}$ . Consider the product  $\hat{f}_n \hat{g}_n^*$ :

$$\begin{aligned} \hat{f}_n \hat{g}_n^* &= |\hat{f}_n| |\hat{g}_n| e^{i(\theta_n - \phi_n)} \\ &= |\hat{f}_n| |\hat{g}_n| [\cos(\theta_n - \phi_n) + i \sin(\theta_n - \phi_n)] \end{aligned} \tag{A.33}$$

In the case that the two quantities are oscillating in phase,  $\theta_n = \phi_n$ , then this product is purely real,  $\text{Im} [\hat{f}_n \hat{g}_n^*] = 0$ . If however there is a phase difference between the quantities  $\theta_n \neq \phi_n$  the imaginary part is non-zero, and thus the imaginary part of the product parameterises the phase difference between the oscillations.

## Appendix B

# Gyrokinetic results

### B.1 Time average

For a given time-dependent function over a domain  $\tau$ , we may express it as a Fourier series,

$$f(t) = \sum_n \hat{f}_n e^{i\omega_n t} \quad (\text{B.1})$$

where  $\omega_n = 2\pi n/\tau$ . We want to integrate it by equation 3.3, for some  $T_0 < \tau$ . We find

$$\begin{aligned} \bar{f}(t) &= \sum_n \hat{f}_n \frac{1}{T_0} \int_{t-T_0/2}^{t+T_0/2} e^{i\omega_n t'} dt' \\ &= \sum_n \hat{f}_n \frac{1}{i\omega_n T_0} e^{i\omega_n t} \left[ e^{i\omega_n T_0/2} - e^{-i\omega_n T_0/2} \right] \\ &= \sum_n \hat{f}_n \text{sinc} \left( \frac{\omega_n T_0}{2} \right) e^{i\omega_n t}. \end{aligned} \quad (\text{B.2})$$

Taking this average again yields

$$\bar{\bar{f}}(t) = \sum_n \hat{f}_n \text{sinc}^2 \left( \frac{\omega_n T_0}{2} \right) e^{i\omega_n t} \neq \bar{f}(t). \quad (\text{B.3})$$

The only case for which  $\bar{\bar{f}} = \bar{f}$  is when  $\omega_n T_0/2$  forms integer multiples of  $\pi$  for all  $n$ , such that  $\omega_n = \frac{2\pi n}{T_0}$ , satisfied when  $T_0 = \tau$ .

### B.2 First order gyromotion contribution

We would like to evaluate  $\langle \mathbf{u}_0 \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}] \rangle_{\mathbf{R}}$ , where gyrophase dependence is present in both  $\mathbf{u}_0$  and  $\boldsymbol{\rho}_0$ . Using equations 3.41 and 3.44, we have  $\mathbf{u}_0 = u_0 (\cos \varphi \hat{\mathbf{e}}_1 - \sin \varphi \hat{\mathbf{e}}_2)$  and  $\boldsymbol{\rho}_0 = \frac{u_0}{\Omega_s} (\sin \varphi \hat{\mathbf{e}}_1 + \cos \varphi \hat{\mathbf{e}}_2)$ . For brevity we label these components  $u_1 = u_0 \cos \varphi$ ,  $u_2 = -u_0 \sin \varphi$ ,  $\rho_1 = (u_0/\Omega_s) \sin \varphi$  and  $\rho_2 = (u_0/\Omega_s) \cos \varphi$ , such that  $\mathbf{u}_0 = u_1 \hat{\mathbf{e}}_2 + u_2 \hat{\mathbf{e}}_1$  and  $\boldsymbol{\rho}_0 = \rho_1 \hat{\mathbf{e}}_1 + \rho_2 \hat{\mathbf{e}}_2$ .

Working in the orthonormal basis  $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{b}}\}$  we first relabel  $(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B} = \mathbf{V} = V_1 \hat{\mathbf{e}}_1 + V_2 \hat{\mathbf{e}}_2 + V_{\parallel} \hat{\mathbf{b}}$  for brevity, where the components are currently unknown. The desired equation becomes

$$\begin{aligned} \langle \mathbf{u}_0 \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}] \rangle_{\mathbf{R}} &= \left\langle (u_1 \hat{\mathbf{e}}_1 + u_2 \hat{\mathbf{e}}_2) \times \left( V_1 \hat{\mathbf{e}}_1 + V_2 \hat{\mathbf{e}}_2 + V_{\parallel} \hat{\mathbf{b}} \right) \right\rangle_{\mathbf{R}} \\ &= \langle u_2 V_{\parallel} \rangle_{\mathbf{R}} \hat{\mathbf{e}}_1 - \langle u_1 V_{\parallel} \rangle_{\mathbf{R}} \hat{\mathbf{e}}_2 + \langle u_1 V_2 - u_2 V_1 \rangle_{\mathbf{R}} \hat{\mathbf{b}}. \end{aligned} \quad (\text{B.4})$$

### $\nabla$ in a local orthonormal basis

To find the components of  $\mathbf{V}$ , we consider  $\nabla$  in this basis. For brevity, we denote the Cartesian basis vectors  $\hat{\mathbf{x}}_j = \{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ , the Cartesian coordinates  $x_j = \{x, y, z\}$ , and the basis vectors in the local orthonormal system  $\hat{\mathbf{e}}_i = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{b}}\}$ . Therefore we may express the basis vectors in each system in terms of the other, via

$$\hat{\mathbf{e}}_i = \sum_j (\hat{\mathbf{e}}_i \cdot \hat{\mathbf{x}}_j) \hat{\mathbf{x}}_j \quad (\text{B.5})$$

$$\hat{\mathbf{x}}_j = \sum_i (\hat{\mathbf{e}}_i \cdot \hat{\mathbf{x}}_j) \hat{\mathbf{e}}_i. \quad (\text{B.6})$$

The object  $\nabla$  is defined

$$\nabla = \sum_j \hat{\mathbf{x}}_j \frac{\partial}{\partial x_j} \quad (\text{B.7})$$

and therefore one may define the partial derivatives in the local system via

$$\frac{\partial}{\partial x_j} = \hat{\mathbf{x}}_j \cdot \nabla = \sum_i (\hat{\mathbf{e}}_i \cdot \hat{\mathbf{x}}_j) (\hat{\mathbf{e}}_i \cdot \nabla). \quad (\text{B.8})$$

Equation B.7 becomes

$$\begin{aligned} \nabla &= \sum_i \sum_j \hat{\mathbf{x}}_j (\hat{\mathbf{e}}_i \cdot \hat{\mathbf{x}}_j) (\hat{\mathbf{e}}_i \cdot \nabla) \\ &= \sum_i \hat{\mathbf{e}}_i (\hat{\mathbf{e}}_i \cdot \nabla) \end{aligned} \quad (\text{B.9})$$

and thus the gradient of a scalar in the local orthonormal system is

$$\nabla f = \hat{\mathbf{e}}_1 \nabla_1 f + \hat{\mathbf{e}}_2 \nabla_2 f + \hat{\mathbf{b}} \nabla_{\parallel} f \quad (\text{B.10})$$

where  $\nabla_1 = \hat{\mathbf{e}}_1 \cdot \nabla$ ,  $\nabla_2 = \hat{\mathbf{e}}_2 \cdot \nabla$  and  $\nabla_{\parallel} = \hat{\mathbf{b}} \cdot \nabla$ .

We next require the divergence of a vector, the definition of which in our notation is

$$\nabla \cdot \mathbf{A} = \sum_j \frac{\partial}{\partial x_j} (\mathbf{A} \cdot \hat{\mathbf{x}}_j). \quad (\text{B.11})$$

Because  $\partial \hat{\mathbf{x}}_j / \partial x_j = \mathbf{0}$ , the basis vector can be taken out of the partial derivative. Then using the above results,

$$\begin{aligned}\nabla \cdot \mathbf{A} &= \sum_j \hat{\mathbf{x}}_j \cdot \frac{\partial \mathbf{A}}{\partial x_j} \\ &= \sum_j \sum_i (\hat{\mathbf{e}}_i \cdot \hat{\mathbf{x}}_j) \hat{\mathbf{x}}_j \cdot [(\hat{\mathbf{e}}_i \cdot \nabla) \mathbf{A}] \\ &= \sum_i \hat{\mathbf{e}}_i \cdot [(\hat{\mathbf{e}}_i \cdot \nabla) \mathbf{A}].\end{aligned}\tag{B.12}$$

The divergence of a vector in the orthonormal basis is therefore

$$\nabla \cdot \mathbf{A} = \hat{\mathbf{e}}_1 \cdot (\nabla_1 \mathbf{A}) + \hat{\mathbf{e}}_2 \cdot (\nabla_2 \mathbf{A}) + \hat{\mathbf{b}} \cdot (\nabla_{\parallel} \mathbf{A}).\tag{B.13}$$

Specifically for the unit vector in the direction of the magnetic field, one has

$$\nabla \cdot \hat{\mathbf{b}} = \hat{\mathbf{e}}_1 \cdot (\nabla_1 \hat{\mathbf{b}}) + \hat{\mathbf{e}}_2 \cdot (\nabla_2 \hat{\mathbf{b}}) + \hat{\mathbf{b}} \cdot (\nabla_{\parallel} \hat{\mathbf{b}})\tag{B.14}$$

where the last term is zero by  $\hat{\mathbf{b}} \cdot (\nabla_{\parallel} \hat{\mathbf{b}}) = \frac{1}{2} \nabla_{\parallel} (\hat{\mathbf{b}} \cdot \hat{\mathbf{b}}) = \frac{1}{2} \nabla_{\parallel} (1) = 0$ .

### Evaluating the gyroaverage

Returning to  $(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}$ , then using the chain rule

$$(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B} = B [(\boldsymbol{\rho}_0 \cdot \nabla) \hat{\mathbf{b}}] + \hat{\mathbf{b}} [(\boldsymbol{\rho}_0 \cdot \nabla) B]\tag{B.15}$$

from which we find  $V_{\parallel} = \rho_1 \nabla_1 B + \rho_2 \nabla_2 B$ . For  $V_1$ , we have  $V_1 = B [(\rho_1 \nabla_1 + \rho_2 \nabla_2) \hat{\mathbf{b}}] \cdot \hat{\mathbf{e}}_1$ , and  $V_2 = B [(\rho_1 \nabla_1 + \rho_2 \nabla_2) \hat{\mathbf{b}}] \cdot \hat{\mathbf{e}}_2$ . Inserting these results back into equation B.4, we get several gyroaverages of the form  $\langle u_i \rho_j \rangle_{\mathbf{R}}$ . Calculated explicitly, these are

$$\langle u_1 \rho_1 \rangle_{\mathbf{R}} = -\langle u_2 \rho_2 \rangle_{\mathbf{R}} = \frac{u_0^2}{\Omega_s} \langle \sin \varphi \cos \varphi \rangle_{\mathbf{R}} = 0\tag{B.16}$$

$$\langle u_1 \rho_2 \rangle_{\mathbf{R}} = \frac{u_0^2}{\Omega_s} \langle \cos^2 \varphi \rangle_{\mathbf{R}} = \frac{u_0^2}{2\Omega_s}\tag{B.17}$$

$$\langle u_2 \rho_1 \rangle_{\mathbf{R}} = -\frac{u_0^2}{\Omega_s} \langle \sin^2 \varphi \rangle_{\mathbf{R}} = -\frac{u_0^2}{2\Omega_s}\tag{B.18}$$

and thus the terms that remain in equation B.4 are

$$\langle \mathbf{u}_0 \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}] \rangle_{\mathbf{R}} = -\frac{u_0^2}{2\Omega_s} \left( \nabla_1 B \hat{\mathbf{e}}_1 + \nabla_2 B \hat{\mathbf{e}}_2 - B \left[ \hat{\mathbf{e}}_1 \cdot (\nabla_1 \hat{\mathbf{b}}) + \hat{\mathbf{e}}_2 \cdot (\nabla_2 \hat{\mathbf{b}}) \right] \hat{\mathbf{b}} \right).\tag{B.19}$$

We recognise the term in the square bracket to be  $\nabla \cdot \hat{\mathbf{b}}$  from equation B.14. Using Gauss's law for magnetism with the chain rule, we have  $\nabla \cdot \mathbf{B} = B \nabla \cdot \hat{\mathbf{b}} + (\hat{\mathbf{b}} \cdot \nabla) B = 0$  and thus

$$\nabla \cdot \hat{\mathbf{b}} = -\frac{1}{B} \nabla_{\parallel} B\tag{B.20}$$

which when inserted into equation B.19, and using the definition of the gradient given by equation B.9, finally gives

$$\langle \mathbf{u}_0 \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}] \rangle_{\mathbf{R}} = -\frac{u_0^2}{2\Omega_s} \nabla B. \quad (\text{B.21})$$

### B.3 Properties of the gyroaverage

We first relate derivatives with respect to particle position to derivatives with respect to gyrocentre position. Using the chain rule and the notation  $\nabla = \partial/\partial\mathbf{r}$ ,

$$\begin{aligned} \frac{\partial}{\partial\mathbf{r}} &= \frac{\partial}{\partial t} + \frac{\partial}{\partial\mathbf{R}} \frac{\partial\mathbf{R}}{\partial\mathbf{r}} + \frac{\partial}{\partial v_1} \frac{\partial v_1}{\partial\mathbf{r}} + \frac{\partial}{\partial v_2} \frac{\partial v_2}{\partial\mathbf{r}} + \frac{\partial}{\partial\varphi} \frac{\partial\varphi}{\partial\mathbf{r}} \\ &= \frac{\partial}{\partial\mathbf{R}} + \mathcal{O}(\rho_* k_\perp) \end{aligned} \quad (\text{B.22})$$

and thus  $\nabla = \partial/\partial\mathbf{R}$  to lowest order for both fluctuating and equilibrium quantities. Taking the gyroaverage of both sides for some function  $\delta g(t, \mathbf{R}, v_1, v_2, \varphi)$  at constant  $\mathbf{R}$ , this commutes only with the  $\mathbf{R}$  derivatives, such that

$$\begin{aligned} \langle \nabla \delta g \rangle_{\mathbf{R}} &= \left\langle \frac{\partial \delta g}{\partial \mathbf{R}} \right\rangle_{\mathbf{R}} + \mathcal{O}(\rho_* k_\perp \delta g) \\ &= \frac{\partial \langle \delta g \rangle_{\mathbf{R}}}{\partial \mathbf{R}} + \mathcal{O}(\rho_* k_\perp \delta g). \end{aligned} \quad (\text{B.23})$$

Another result of the gyroaverage will now be shown, where for clarity we make the partial derivatives explicit. Consider phase coordinates  $\{t, \mathbf{r}, v_1, v_2, \varphi\}$ . Writing the derivative with respect to gyrophase at constant  $\mathbf{R}$  in this system via the chain rule we have

$$\left. \frac{\partial \delta g}{\partial \varphi} \right|_{t, \mathbf{R}, v_1, v_2} = \left. \frac{\partial \delta g}{\partial \mathbf{r}} \right|_{t, v_1, v_2, \varphi} \cdot \left. \frac{\partial \mathbf{r}}{\partial \varphi} \right|_{t, \mathbf{R}, v_1, v_2} + \left. \frac{\partial \delta g}{\partial \varphi} \right|_{t, \mathbf{r}, v_1, v_2} \quad (\text{B.24})$$

where  $\left. \frac{\partial \delta g}{\partial \mathbf{r}} \right|_{t, v_1, v_2, \varphi} = \nabla \delta g$  and

$$\left. \frac{\partial \mathbf{r}}{\partial \varphi} \right|_{t, \mathbf{R}, v_1, v_2} = \frac{1}{\Omega_s} \mathbf{u}_0 + \mathcal{O}(\rho_* \rho) \quad (\text{B.25})$$

where we have used 3.14 and 3.43. Equation B.24 becomes

$$\left. \frac{\partial \delta g}{\partial \varphi} \right|_{t, \mathbf{R}, v_1, v_2} = \frac{1}{\Omega_s} \mathbf{u}_0 \cdot \nabla_\perp \delta g + \left. \frac{\partial \delta g}{\partial \varphi} \right|_{t, \mathbf{r}, v_1, v_2} + \mathcal{O}(\rho_* \delta g). \quad (\text{B.26})$$

which upon gyroaveraging both sides at constant  $\mathbf{R}$  yields

$$\frac{1}{\Omega_s} \langle \mathbf{u}_0 \cdot \nabla_\perp \delta g \rangle_{\mathbf{R}} = - \left\langle \left. \frac{\partial \delta g}{\partial \varphi} \right|_{t, \mathbf{r}, v_1, v_2} \right\rangle_{\mathbf{R}} + \mathcal{O}(\rho_* \delta g). \quad (\text{B.27})$$

Considering the electrostatic potential, which is independent of  $\varphi$  at constant  $\mathbf{r}$ ,  $\delta\phi_1 = \delta\phi_1(\mathbf{r}, t)$ , one obtains

$$\langle \mathbf{u}_0 \cdot \nabla_\perp \delta\phi_1 \rangle_{\mathbf{R}} = \mathcal{O}(\rho_* \Omega_s \delta\phi_1) \quad (\text{B.28})$$

and so vanishes to lowest order.

## B.4 Velocity coordinate evolution equations

Starting from equation 3.53 and using equations 3.41 and 3.34 one obtains

$$\left(\frac{dU_{0,\parallel}}{dt}\right)_1 \hat{\mathbf{b}} + U_{0,\parallel} \left(\frac{d\hat{\mathbf{b}}}{dt}\right)_1 + \left(\frac{du_0}{dt}\right)_1 \hat{\mathbf{e}}_\perp + u_0 \left(\frac{d\hat{\mathbf{e}}_\perp}{dt}\right)_1 + \frac{\partial \mathbf{u}_1}{\partial \varphi} \Omega_s = \frac{1}{m_s} \mathbf{F}_1. \quad (\text{B.29})$$

We can isolate the evolution equations of  $U_{0,\parallel}$  and  $u_0$  by taking the dot product of this equation with different basis vectors.

### Parallel velocity evolution

Taking the dot product of equation B.29 with  $\hat{\mathbf{b}}$ , one finds

$$\left(\frac{dU_{0,\parallel}}{dt}\right)_1 - u_0 \left(\frac{d\hat{\mathbf{b}}}{dt}\right)_1 \cdot \hat{\mathbf{e}}_\perp + \frac{\partial (\mathbf{u}_1 \cdot \hat{\mathbf{b}})}{\partial \varphi} \Omega_s = \frac{1}{m_s} \mathbf{F}_1 \cdot \hat{\mathbf{b}}. \quad (\text{B.30})$$

where we have used  $d(\hat{\mathbf{b}} \cdot \hat{\mathbf{e}}_\perp)/dt = \hat{\mathbf{b}} \cdot d\hat{\mathbf{e}}_\perp/dt + \hat{\mathbf{e}}_\perp \cdot d\hat{\mathbf{b}}/dt = 0$  and  $\hat{\mathbf{b}} \cdot d\hat{\mathbf{b}}/dt = 0$ . Gyroaveraging both sides at constant  $\mathbf{R}$ ,

$$\left\langle \left(\frac{dU_{0,\parallel}}{dt}\right)_1 \right\rangle_{\mathbf{R}} = \frac{1}{m_s} \langle \mathbf{F}_1 \rangle_{\mathbf{R}} \cdot \hat{\mathbf{b}} \quad (\text{B.31})$$

which, using 3.55 and 3.57, becomes

$$\left\langle \left(\frac{dU_{0,\parallel}}{dt}\right)_1 \right\rangle_{\mathbf{R}} = -\frac{u_0^2}{2B} \hat{\mathbf{b}} \cdot \nabla B. \quad (\text{B.32})$$

### Perpendicular velocity evolution

Now taking the dot product of equation B.29 with  $\hat{\mathbf{e}}_\perp$ ,

$$U_{0,\parallel} \left(\frac{d\hat{\mathbf{b}}}{dt}\right)_1 \cdot \hat{\mathbf{e}}_\perp + \left(\frac{du_0}{dt}\right)_1 + \frac{\partial \mathbf{u}_1}{\partial \varphi} \cdot \hat{\mathbf{e}}_\perp \Omega_s = \frac{1}{m_s} \mathbf{F}_1 \cdot \hat{\mathbf{e}}_\perp. \quad (\text{B.33})$$

Gyroaveraging and rearranging, we get

$$\left\langle \left(\frac{du_0}{dt}\right)_1 \right\rangle_{\mathbf{R}} = \frac{1}{m_s} \langle \mathbf{F}_1 \cdot \hat{\mathbf{e}}_\perp \rangle_{\mathbf{R}} - \left\langle \frac{\partial \mathbf{u}_1}{\partial \varphi} \cdot \hat{\mathbf{e}}_\perp \right\rangle_{\mathbf{R}} \Omega_s. \quad (\text{B.34})$$

Considering the force term, using 3.32 we have

$$\frac{1}{m_s} \langle \mathbf{F}_1 \cdot \hat{\mathbf{e}}_\perp \rangle_{\mathbf{R}} = \frac{q_s}{m_s} [\langle \hat{\mathbf{e}}_\perp \cdot (\mathbf{u}_1 \times \mathbf{B}) \rangle_{\mathbf{R}} + \langle \hat{\mathbf{e}}_\perp \cdot (\mathbf{U}_0 \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}]) \rangle_{\mathbf{R}}] \quad (\text{B.35})$$

where the potential term is zero due to B.28. For the first term in equation B.35, we use the triple scalar product to write  $\hat{\mathbf{e}}_\perp \cdot (\mathbf{u}_1 \times \mathbf{B}) = -B \mathbf{u}_1 \cdot (\hat{\mathbf{e}}_\perp \times \hat{\mathbf{b}})$ . Using the triple scalar product again on the second term, one has

$$\begin{aligned} \langle \hat{\mathbf{e}}_\perp \cdot (\mathbf{U}_0 \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}]) \rangle_{\mathbf{R}} &= -\mathbf{U}_0 \cdot \langle \hat{\mathbf{e}}_\perp \times [(\boldsymbol{\rho}_0 \cdot \nabla) \mathbf{B}] \rangle_{\mathbf{R}} \\ &= \frac{u_0}{2\Omega_s} U_{0,\parallel} \hat{\mathbf{b}} \cdot \nabla B \end{aligned} \quad (\text{B.36})$$

where we have used 3.57. Equation B.34 therefore becomes

$$\left\langle \left( \frac{du_0}{dt} \right)_1 \right\rangle_{\mathbf{R}} = \frac{u_0}{2B} U_{0,\parallel} \hat{\mathbf{b}} \cdot \nabla B_0 - \left[ \left\langle \frac{\partial \mathbf{u}_1}{\partial \varphi} \cdot \hat{\mathbf{e}}_{\perp} \right\rangle_{\mathbf{R}} + \left\langle \mathbf{u}_1 \cdot (\hat{\mathbf{e}}_{\perp} \times \hat{\mathbf{b}}) \right\rangle_{\mathbf{R}} \right] \Omega_s. \quad (\text{B.37})$$

By noting that  $\hat{\mathbf{e}}_{\perp} \times \hat{\mathbf{b}} = \partial \hat{\mathbf{e}}_{\perp} / \partial \varphi$ , the square bracket is shown to vanish via

$$\left\langle \frac{\partial \mathbf{u}_1}{\partial \varphi} \cdot \hat{\mathbf{e}}_{\perp} \right\rangle_{\mathbf{R}} + \left\langle \mathbf{u}_1 \cdot \frac{\partial \hat{\mathbf{e}}_{\perp}}{\partial \varphi} \right\rangle_{\mathbf{R}} = \left\langle \frac{\partial (\mathbf{u}_1 \cdot \hat{\mathbf{e}}_{\perp})}{\partial \varphi} \right\rangle_{\mathbf{R}} = 0 \quad (\text{B.38})$$

and thus we obtain

$$\left\langle \left( \frac{du_0}{dt} \right)_1 \right\rangle_{\mathbf{R}} = \frac{u_0}{2B} U_{0,\parallel} \hat{\mathbf{b}} \cdot \nabla B. \quad (\text{B.39})$$

## B.5 Derivation of the gyrokinetic equation

The distribution function of species  $s$  up to second order is

$$f_s = F_{s,0} + F_{s,1} + \delta f_{s,1,A} + h_s + F_{s,2} + \delta f_2 + \mathcal{O}(\rho_*^3 F_{s,0}) \quad (\text{B.40})$$

where  $\delta f_{s,1,A} = -q_s \delta \phi F_{s,0} / T_{s,0}$  is the adiabatic part of the first order fluctuating distribution and  $F_{s,0}$  is the Maxwellian background given by equation 3.92. We shall derive the gyrokinetic equation by returning to the original statement of the kinetic equation

$$\frac{df_s}{dt} = C_s \quad (\text{B.41})$$

and using equation B.40. The gyrokinetic equation is the fluctuating part of the kinetic equation for terms of order  $\mathcal{O}(\rho_*^2 \Omega_s F_{s,0})$ , gyroaveraged at constant  $\mathbf{R}$ . Applying these operations to equation B.41, we have

$$\begin{aligned} \langle (\delta C_s)_2 \rangle_{\mathbf{R}} = & \left\langle \delta \left( \frac{dF_{s,0}}{dt} \right)_2 \right\rangle_{\mathbf{R}} + \left\langle \delta \left( \frac{dF_{s,1}}{dt} \right)_2 \right\rangle_{\mathbf{R}} + \left\langle \delta \left( \frac{dF_{s,2}}{dt} \right)_2 \right\rangle_{\mathbf{R}} \\ & + \left\langle \delta \left( \frac{d\delta f_{s,1,A}}{dt} \right)_2 \right\rangle_{\mathbf{R}} + \left\langle \delta \left( \frac{dh_s}{dt} \right)_2 \right\rangle_{\mathbf{R}} + \left\langle \delta \left( \frac{d\delta f_{s,2}}{dt} \right)_2 \right\rangle_{\mathbf{R}}. \end{aligned} \quad (\text{B.42})$$

Each term shall now be considered individually, working with the phase coordinates  $\{t, \mathbf{R}, \varepsilon, \mu, \varphi\}$ . Note, taking the fluctuating component commutes with taking the gyroaverage, however neither of these commute with taking the *total* time derivative, only with taking partial derivatives.

### The terms of the equation B.42

Starting with the  $F_{s,2}$  term, we have

$$\left( \frac{dF_{s,2}}{dt} \right)_2 = \Omega_s \frac{\partial F_{s,2}}{\partial \varphi} \Big|_{t, \mathbf{R}, \varepsilon, \mu} \implies \left\langle \left( \frac{dF_{s,2}}{dt} \right)_2 \right\rangle_{\mathbf{R}} = 0 \quad (\text{B.43})$$

and similarly for the  $\delta f_{s,2}$  term,

$$\left(\frac{d\delta f_{s,2}}{dt}\right)_2 = \Omega_s \frac{\partial \delta f_{s,2}}{\partial \varphi} \Big|_{t,\mathbf{R},\epsilon,\mu} \implies \left\langle \left(\frac{d\delta f_{s,2}}{dt}\right)_2 \right\rangle_{\mathbf{R}} = 0. \quad (\text{B.44})$$

The second order parts of the distribution function are therefore not present in the gyrokinetic equation. Now turning to the  $F_{s,1}$  term, using  $\partial F_{s,1}/\partial \varphi = 0$  from equation 3.94, we have

$$\left(\frac{dF_{s,1}}{dt}\right)_2 = \mathbf{U}_0 \cdot \frac{\partial F_{s,1}}{\partial \mathbf{R}} \Big|_{t,\epsilon,\mu,\varphi} + \left(\frac{d\epsilon}{dt}\right)_1 \frac{\partial F_{s,1}}{\partial \epsilon} \Big|_{t,\mathbf{R},\mu,\varphi} + \left(\frac{d\mu}{dt}\right)_1 \frac{\partial F_{s,1}}{\partial \mu} \Big|_{t,\mathbf{R},\epsilon,\varphi}. \quad (\text{B.45})$$

Applying the gyroaverage and using equations 3.68 and 3.72 yields

$$\left\langle \left(\frac{dF_{s,1}}{dt}\right)_2 \right\rangle_{\mathbf{R}} = \mathbf{U}_0 \cdot \frac{\partial F_{s,1}}{\partial \mathbf{R}} \Big|_{t,\epsilon,\mu,\varphi}. \quad (\text{B.46})$$

Taking the fluctuating component gives

$$\delta \left\langle \left(\frac{dF_{s,1}}{dt}\right)_2 \right\rangle_{\mathbf{R}} = 0 \quad (\text{B.47})$$

and therefore  $F_{s,1}$  does not appear in the gyrokinetic equation either.

Turning now to the  $F_{s,0}$  term,

$$\left(\frac{dF_{s,0}}{dt}\right)_2 = \left(\frac{d\mathbf{R}}{dt}\right)_1 \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} \Big|_{t,\epsilon,\mu,\varphi} + \left(\frac{d\epsilon}{dt}\right)_2 \frac{\partial F_{s,0}}{\partial \epsilon} \Big|_{t,\mathbf{R},\mu,\varphi} \quad (\text{B.48})$$

we see the presence of energy evolution to second order. This does not require explicit evaluation as it will be shown to cancel with an upcoming term. Taking the fluctuating component of equation B.48, using equation 3.60, evaluating the energy derivative and gyroaveraging we get

$$\left\langle \delta \left(\frac{dF_{s,0}}{dt}\right)_2 \right\rangle_{\mathbf{R}} = \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} \Big|_{t,\epsilon,\mu,\varphi} - \left\langle \delta \left(\frac{d\epsilon}{dt}\right)_2 \right\rangle_{\mathbf{R}} \frac{F_{s,0}}{T_{s,0}}. \quad (\text{B.49})$$

Now considering the  $h_s$  term, where  $h_s$  is independent of gyrophase from equation 3.96, we have

$$\left(\frac{dh_s}{dt}\right)_2 = \frac{\partial h_s}{\partial t} + \left(\frac{d\mathbf{R}}{dt}\right)_2 \cdot \frac{\partial h_s}{\partial \mathbf{R}} + \left(\frac{d\epsilon}{dt}\right)_1 \frac{\partial h_s}{\partial \epsilon} + \left(\frac{d\mu}{dt}\right)_1 \frac{\partial h_s}{\partial \mu}. \quad (\text{B.50})$$

Gyroaveraging this equation yields

$$\left\langle \left(\frac{dh_s}{dt}\right)_2 \right\rangle_{\mathbf{R}} = \frac{\partial h_s}{\partial t} + \left\langle \left(\frac{d\mathbf{R}}{dt}\right)_2 \right\rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}}. \quad (\text{B.51})$$

Evaluating the second term using equation 3.60 gives

$$\left\langle \left(\frac{d\mathbf{R}}{dt}\right)_2 \right\rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} = \left[ U_{0,\parallel} \hat{\mathbf{b}} + \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} + \mathbf{v}_D \right] \cdot \frac{\partial h_s}{\partial \mathbf{R}}. \quad (\text{B.52})$$

Note that the first order parallel velocity  $U_{1,\parallel}$  is not present here as  $U_{1,\parallel} \hat{\mathbf{b}} \cdot \partial h_s / \partial \mathbf{R} = \mathcal{O}(\rho_*^3 \Omega_s F_{s,0})$ . The ensemble-averaged part of equation B.51 is therefore

$$\left\langle \left\langle \left(\frac{dh_s}{dt}\right)_2 \right\rangle_{\mathbf{R}} \right\rangle_{\text{Ens}} = \left\langle \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} \right\rangle_{\text{Ens}} \quad (\text{B.53})$$

giving the fluctuating component as

$$\begin{aligned} \delta \left\langle \left( \frac{dh_s}{dt} \right)_2 \right\rangle_{\mathbf{R}} &= \left\langle \left( \frac{dh_s}{dt} \right)_2 \right\rangle_{\mathbf{R}} - \left\langle \left\langle \left( \frac{dh_s}{dt} \right)_2 \right\rangle_{\mathbf{R}} \right\rangle_{\text{Ens}} \\ &= \frac{\partial h_s}{\partial t} + [U_{0,\parallel} \hat{\mathbf{b}} + \mathbf{v}_D] \cdot \frac{\partial h_s}{\partial \mathbf{R}} + \delta \left( \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} \right) \end{aligned} \quad (\text{B.54})$$

where

$$\delta \left( \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} \right) = \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} - \left\langle \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} \right\rangle_{\text{Ens}}. \quad (\text{B.55})$$

Now turning our attention to the term involving the adiabatic part of the first order fluctuating distribution function, we find

$$\begin{aligned} \left( \frac{d\delta f_{s,1,A}}{dt} \right)_2 &= -\frac{q_s}{T_{s,0}} F_{s,0} \left( \frac{d\delta\phi_1}{dt} \right)_2 - q_s \delta\phi_1 \left( \frac{d}{dt} \left[ \frac{F_{s,0}}{T_{s,0}} \right] \right)_1 \\ &= -\frac{q_s}{T_{s,0}} F_{s,0} \left( \frac{d\delta\phi_1}{dt} \right)_2 - q_s \delta\phi_1 \left[ \mathbf{U}_0 \cdot \frac{\partial}{\partial \mathbf{R}} \left( \frac{F_{s,0}}{T_{s,0}} \right) - \left( \frac{d\varepsilon}{dt} \right)_1 \frac{F_{s,0}}{T_{s,0}^2} \right] \end{aligned} \quad (\text{B.56})$$

where the derivatives parallel to the magnetic field are zero from equation 3.91. Gyroaveraging the remaining terms gives

$$\left\langle \left( \frac{d\delta f_{s,1,A}}{dt} \right)_2 \right\rangle_{\mathbf{R}} = -\frac{q_s}{T_{s,0}} F_{s,0} \left\langle \left( \frac{d\delta\phi_1}{dt} \right)_2 \right\rangle_{\mathbf{R}} + q_s \frac{F_{s,0}}{T_{s,0}^2} \left\langle \delta\phi_1 \left( \frac{d\varepsilon}{dt} \right)_1 \right\rangle_{\mathbf{R}} \quad (\text{B.57})$$

for which the last term is zero via

$$\begin{aligned} \left\langle \delta\phi_1 \left( \frac{d\varepsilon}{dt} \right)_1 \right\rangle_{\mathbf{R}} &= -q_s \langle \delta\phi_1 (\mathbf{u}_0 \cdot \nabla_{\perp} \delta\phi_1) \rangle_{\mathbf{R}} \\ &= -\frac{1}{2} q_s \langle \mathbf{u}_0 \cdot \nabla_{\perp} (\delta\phi_1)^2 \rangle_{\mathbf{R}} \\ &= 0 \end{aligned} \quad (\text{B.58})$$

where we have used equation B.27. The fluctuating part of equation B.57 is therefore

$$\delta \left\langle \left( \frac{d\delta f_{s,1,A}}{dt} \right)_2 \right\rangle_{\mathbf{R}} = -\frac{q_s}{T_{s,0}} F_{s,0} \left\langle \delta \left( \frac{d\delta\phi_1}{dt} \right)_2 \right\rangle_{\mathbf{R}}. \quad (\text{B.59})$$

### Bringing the terms together

Now writing equation B.42 using the results obtained thus far, we have

$$\begin{aligned} \frac{\partial h_s}{\partial t} + [U_{0,\parallel} \hat{\mathbf{b}} + \mathbf{v}_D] \cdot \frac{\partial h_s}{\partial \mathbf{R}} + \delta \left( \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} \right) - \frac{q_s}{T_{s,0}} F_{s,0} \left\langle \delta \left( \frac{d\delta\phi_1}{dt} \right)_2 \right\rangle_{\mathbf{R}} + \\ \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} \Big|_{t,\epsilon,\mu,\varphi} - \left\langle \delta \left( \frac{d\varepsilon}{dt} \right)_2 \right\rangle_{\mathbf{R}} \frac{F_{s,0}}{T_{s,0}} = \langle (\delta C_s)_2 \rangle_{\mathbf{R}}. \end{aligned} \quad (\text{B.60})$$

The final simplification can be made by considering conservation of energy, which from equation 3.69 is

$$\begin{aligned}\frac{d\varepsilon}{dt} &= -q_s \mathbf{v} \cdot \nabla \delta\phi \\ &= q_s \left. \frac{\partial \delta\phi}{\partial t} \right|_{\mathbf{r}, \mathbf{v}} - q_s \frac{d\delta\phi}{dt}.\end{aligned}\tag{B.61}$$

Taking the fluctuating component and the gyroaverage of this equation, the second order terms are

$$\begin{aligned}\left\langle \delta \left( \frac{d\varepsilon}{dt} \right)_2 \right\rangle_{\mathbf{R}} + q_s \left\langle \delta \left( \frac{d\delta\phi}{dt} \right)_2 \right\rangle_{\mathbf{R}} &= q_s \left\langle \left. \frac{\partial \delta\phi_1}{\partial t} \right|_{\mathbf{r}, \mathbf{v}} \right\rangle_{\mathbf{R}} \\ &= q_s \left. \frac{\partial \langle \delta\phi_1 \rangle_{\mathbf{R}}}{\partial t} \right|_{\mathbf{R}, \varepsilon, \mu, \varphi}.\end{aligned}\tag{B.62}$$

The gyrokinetic equation can therefore finally be written

$$\frac{\partial h_s}{\partial t} + [U_{0,\parallel} \hat{\mathbf{b}} + \mathbf{v}_D] \cdot \frac{\partial h_s}{\partial \mathbf{R}} + \delta \left( \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} \right) = \frac{q_s F_{s,0}}{T_{s,0}} \frac{\partial \langle \delta\phi_1 \rangle_{\mathbf{R}}}{\partial t} - \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} + \langle (\delta C_s)_2 \rangle_{\mathbf{R}}\tag{B.63}$$

which is equation 3.97.

## B.6 The gyrokinetic equation in general velocity coordinates

We wish to express the nonlinear gyrokinetic equation in arbitrary velocity coordinates  $v_1$  and  $v_2$ . For clarity we prime the new system, such that  $\{t, \mathbf{R}, \varepsilon, \mu, \varphi\} \rightarrow \{t', \mathbf{R}', v'_1, v'_2, \varphi'\}$ , with inverse transformations

$$t' = t\tag{B.64}$$

$$\mathbf{R}' = \mathbf{R}\tag{B.65}$$

$$v'_1 = v'_1(t, \mathbf{R}, \varepsilon, \mu)\tag{B.66}$$

$$v'_2 = v'_2(t, \mathbf{R}, \varepsilon, \mu)\tag{B.67}$$

$$\varphi' = \varphi\tag{B.68}$$

where neither  $v'_1$  or  $v'_2$  depend on  $\varphi$  and their time and position dependence comes from equilibrium quantities, so as to satisfy equation 3.20, such that  $\partial v'_{1,2}/\partial t = \mathcal{O}(\rho_*^3 \Omega_s v'_{1,2})$  and  $\partial v'_{1,2}/\partial \mathbf{R} = \mathcal{O}(v'_{1,2}/a)$ .

There are two partial derivatives present in equation 3.97, one with respect to time and one with respect to gyrocentre. Acting on some fluctuating quantity  $\delta g$ , the partial derivative

with respect to time transforms as

$$\begin{aligned}\frac{\partial \delta g}{\partial t} \Big|_{\mathbf{R}, \varepsilon, \mu, \varphi} &= \frac{\partial \delta g}{\partial t'} \Big|_{\mathbf{R}', v'_1, v'_2, \varphi'} + \frac{\partial v'_1}{\partial t} \Big|_{\mathbf{R}, \varepsilon, \mu, \varphi} \frac{\partial \delta g}{\partial v'_1} \Big|_{t', \mathbf{R}', v'_2, \varphi'} + \frac{\partial v'_2}{\partial t} \Big|_{\mathbf{R}, \varepsilon, \mu, \varphi} \frac{\partial \delta g}{\partial v'_2} \Big|_{t', \mathbf{R}', v'_1, \varphi'} \\ &= \frac{\partial \delta g}{\partial t'} \Big|_{\mathbf{R}', v'_1, v'_2, \varphi'} + \mathcal{O}(\rho_*^3 \Omega_s \delta g)\end{aligned}\tag{B.69}$$

due to the equilibrium time variation in the velocity coordinates.

For the spatial derivatives, we will first consider the fluctuating quantities in 3.97 before considering the Maxwellian. The partial derivative acting on  $\delta g$  transforms as

$$\frac{\partial \delta g}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} = \underbrace{\frac{\partial \delta g}{\partial \mathbf{R}'} \Big|_{t', v'_1, v'_2, \varphi'}}_{\mathcal{O}(k\delta g)} + \underbrace{\frac{\partial v'_1}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} \frac{\partial \delta g}{\partial v'_1} \Big|_{t', \mathbf{R}', v'_2, \varphi'} + \frac{\partial v'_2}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} \frac{\partial \delta g}{\partial v'_2} \Big|_{t', \mathbf{R}', v'_1, \varphi'}}_{\mathcal{O}(\delta g/a)}\tag{B.70}$$

where the size of the first term on the right-hand side depends on whether one is considering parallel or perpendicular to the magnetic field. Hence for the equilibrium velocities term in equation 3.97, we have

$$\begin{aligned}\left[ U_{0, \parallel} \hat{\mathbf{b}} + \mathbf{v}_D \right] \cdot \frac{\partial h_s}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} &= \left[ U_{0, \parallel} \hat{\mathbf{b}} + \mathbf{v}_D \right] \cdot \frac{\partial h_s}{\partial \mathbf{R}'} \Big|_{t', v'_1, v'_2, \varphi'} \\ &+ U_{0, \parallel} \hat{\mathbf{b}} \cdot \left[ \frac{\partial v'_1}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} \frac{\partial h_s}{\partial v'_1} \Big|_{t', \mathbf{R}', v'_2, \varphi'} + \frac{\partial v'_2}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} \frac{\partial h_s}{\partial v'_2} \Big|_{t', \mathbf{R}', v'_1, \varphi'} \right] \\ &+ \mathcal{O}(\rho_*^3 \Omega_s F_{s,0})\end{aligned}\tag{B.71}$$

and so the derivatives of  $h_s$  with  $v'_1$  and  $v'_2$  are retained.

For the nonlinear term, then using the triple scalar product, we may write

$$\begin{aligned}\langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} &= \frac{\hat{\mathbf{b}}}{B} \cdot \left( \frac{\partial \langle \delta \phi_1 \rangle_{\mathbf{R}}}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} \times \frac{\partial h_s}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} \right) \\ &= \frac{\hat{\mathbf{b}}}{B} \cdot \left( \frac{\partial \langle \delta \phi_1 \rangle_{\mathbf{R}}}{\partial \mathbf{R}'} \Big|_{t', v'_1, v'_2, \varphi'} \times \frac{\partial h_s}{\partial \mathbf{R}'} \Big|_{t', v'_1, v'_2, \varphi'} \right) + \mathcal{O}(\rho_*^3 \Omega_s F_{s,0}) \\ &= \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}'} \Big|_{t', v'_1, v'_2, \varphi'} + \mathcal{O}(\rho_*^3 \Omega_s F_{s,0})\end{aligned}\tag{B.72}$$

where all the remaining terms are small due to the equilibrium spatial variation in  $v'_1$  and  $v'_2$ . The form of the nonlinear term therefore remains unchanged.

Finally we consider the spatial derivative of the Maxwellian in equation 3.97. Evaluating this explicitly using equation 3.92, one obtains

$$\frac{\partial F_{s,0}}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} = F_{s,0} \left[ \frac{1}{n_{s,0}} \frac{\partial n_{s,0}}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} + \frac{1}{T_{s,0}} \frac{\partial T_{s,0}}{\partial \mathbf{R}} \Big|_{t, \varepsilon, \mu, \varphi} \left( \frac{\varepsilon}{T_{s,0}} - \frac{3}{2} \right) \right].\tag{B.73}$$

To calculate this quantity in our general velocity coordinates, then under the transformation the energy becomes a function of the new coordinates,  $\varepsilon = \varepsilon(t', \mathbf{R}', v'_1, v'_2)$ , such that the Maxwellian is written

$$F_{s,0}(t', \mathbf{R}', v'_1, v'_2) = n_{s,0}(t', \mathbf{R}') \left( \frac{m_s}{2\pi T_{s,0}(t', \mathbf{R}')} \right)^{3/2} e^{-\varepsilon(t', \mathbf{R}', v'_1, v'_2)/T_{s,0}(t', \mathbf{R}')}. \quad (\text{B.74})$$

Evaluating the partial derivative using the chain rule, one has

$$\frac{\partial F_{s,0}}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} = \frac{\partial F_{s,0}}{\partial \mathbf{R}'} \Big|_{t',v'_1,v'_2,\varphi'} + \frac{\partial v'_1}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} \frac{\partial F_{s,0}}{\partial v'_1} \Big|_{t',\mathbf{R}',v'_2,\varphi'} + \frac{\partial v'_2}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} \frac{\partial F_{s,0}}{\partial v'_2} \Big|_{t',\mathbf{R}',v'_1,\varphi'} \quad (\text{B.75})$$

where all terms are order  $\mathcal{O}(F_{s,0}/a)$ . This becomes

$$\begin{aligned} \frac{\partial F_{s,0}}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} &= F_{s,0} \left[ \frac{1}{n_{s,0}} \frac{\partial n_{s,0}}{\partial \mathbf{R}'} \Big|_{t',v'_1,v'_2,\varphi'} + \frac{1}{T_{s,0}} \frac{\partial T_{s,0}}{\partial \mathbf{R}'} \Big|_{t',v'_1,v'_2,\varphi'} \left( \frac{\varepsilon}{T_{s,0}} - \frac{3}{2} \right) \right] \\ &\quad - \frac{F_{s,0}}{T_{s,0}} \left[ \frac{\partial \varepsilon}{\partial \mathbf{R}'} \Big|_{t',v'_1,v'_2,\varphi'} + \frac{\partial \varepsilon}{\partial v'_1} \Big|_{t',\mathbf{R}',v'_2,\varphi'} \frac{\partial v'_1}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} + \frac{\partial \varepsilon}{\partial v'_2} \Big|_{t',\mathbf{R}',v'_1,\varphi'} \frac{\partial v'_2}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} \right] \end{aligned} \quad (\text{B.76})$$

where using the chain rule, the second square bracket can be seen to be the partial derivative of the energy  $\varepsilon$  with respect to the gyrocentre *at constant*  $\varepsilon$ ,

$$\frac{\partial \varepsilon}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} = \frac{\partial \varepsilon}{\partial \mathbf{R}'} \Big|_{t',v'_1,v'_2,\varphi'} + \frac{\partial \varepsilon}{\partial v'_1} \Big|_{t',\mathbf{R}',v'_2,\varphi'} \frac{\partial v'_1}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} + \frac{\partial \varepsilon}{\partial v'_2} \Big|_{t',\mathbf{R}',v'_1,\varphi'} \frac{\partial v'_2}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} = \mathbf{0} \quad (\text{B.77})$$

and thus is zero. Therefore the derivative of the Maxwellian written in the new system is simply

$$\frac{\partial F_{s,0}}{\partial \mathbf{R}} \Big|_{t,\varepsilon,\mu,\varphi} = F_{s,0} \left[ \frac{1}{n_{s,0}} \frac{\partial n_{s,0}}{\partial \mathbf{R}'} \Big|_{t',v'_1,v'_2,\varphi'} + \frac{1}{T_{s,0}} \frac{\partial T_{s,0}}{\partial \mathbf{R}'} \Big|_{t',v'_1,v'_2,\varphi'} \left( \frac{\varepsilon(t', \mathbf{R}', v'_1, v'_2)}{T_{s,0}} - \frac{3}{2} \right) \right] \quad (\text{B.78})$$

Bringing the results of this section together, the gyrokinetic equation in the phase coordinates  $\{t', \mathbf{R}', v'_1, v'_2, \varphi'\}$  is

$$\begin{aligned} \frac{\partial h_s}{\partial t'} + [U_{0,\parallel} \hat{\mathbf{b}} + \mathbf{v}_D] \cdot \frac{\partial h_s}{\partial \mathbf{R}'} + \delta \left( \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}'} \right) + U_{0,\parallel} \hat{\mathbf{b}} \cdot \left[ \frac{\partial v'_1}{\partial \mathbf{R}} \frac{\partial h_s}{\partial v'_1} + \frac{\partial v'_2}{\partial \mathbf{R}} \frac{\partial h_s}{\partial v'_2} \right] = \\ \frac{q_s F_{s,0}}{T_{s,0}} \frac{\partial \langle \delta \phi_1 \rangle_{\mathbf{R}}}{\partial t'} - \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} + \langle (\delta C_s)_2 \rangle_{\mathbf{R}} \end{aligned} \quad (\text{B.79})$$

where the Maxwellian derivative is given by equation B.78. The main difference between this formulation and equation 3.97 is the addition of  $v'_1$  and  $v'_2$  derivatives acting on  $h_s$ . We now consider examples of  $v'_1$  and  $v'_2$  from gyrokinetic codes.

### GKW velocity coordinates

The code GKW [35] uses the parallel velocity  $v'_1 = U'_{0,\parallel}$  (written  $v_{\parallel}$  in the GKW literature) and the magnetic moment  $v'_2 = \mu'$  for its velocity coordinates, with transformations  $U'_{0,\parallel} = \sigma \sqrt{\frac{2}{m_s} (\varepsilon - \mu B)}$  and  $\mu' = \mu$ , where  $\sigma$  is a binary coordinate specifying the sign of the parallel velocity. The partial derivatives with respect to  $\mathbf{R}$  used in equation B.79 are therefore

$$\left. \frac{\partial \mu'}{\partial \mathbf{R}} \right|_{t,\varepsilon,\mu,\varphi} = \mathbf{0} \quad (\text{B.80})$$

$$\left. \frac{\partial U'_{0,\parallel}}{\partial \mathbf{R}} \right|_{t,\varepsilon,\mu,\varphi} = -\frac{1}{U'_{0,\parallel}} \frac{\mu}{m_s} \nabla B \quad (\text{B.81})$$

where we have used B.22. Thus the gyrokinetic equation in the GKW coordinates has the term

$$U_{0,\parallel} \hat{\mathbf{b}} \cdot \frac{\partial U'_{0,\parallel}}{\partial \mathbf{R}} \frac{\partial h_s}{\partial U'_{0,\parallel}} = -\frac{\mu}{m_s} \left( \hat{\mathbf{b}} \cdot \nabla B \right) \frac{\partial h_s}{\partial U'_{0,\parallel}}. \quad (\text{B.82})$$

### CGYRO velocity coordinates

CGYRO [38] uses the speed  $v' = \sqrt{2\varepsilon/m_s}$  and the ‘pitch angle’, defined as the ratio of the parallel velocity to the speed,  $\xi' = \sigma \sqrt{\frac{2}{m_s} (\varepsilon - \mu B)} / \sqrt{2\varepsilon/m_s} = \sigma \sqrt{1 - \mu B/\varepsilon}$ . The partial derivatives are therefore

$$\left. \frac{\partial v'}{\partial \mathbf{R}} \right|_{t,\varepsilon,\mu,\varphi} = \mathbf{0} \quad (\text{B.83})$$

$$\left. \frac{\partial \xi'}{\partial \mathbf{R}} \right|_{t,\varepsilon,\mu,\varphi} = -\frac{1}{2\xi' B} \left( 1 - (\xi')^2 \right) \nabla B \quad (\text{B.84})$$

and thus the gyrokinetic equation in CGYRO has the term

$$U_{0,\parallel} \hat{\mathbf{b}} \cdot \frac{\partial \xi'}{\partial \mathbf{R}} \frac{\partial h_s}{\partial \xi'} = -\frac{v'}{2B} \left( 1 - (\xi')^2 \right) \hat{\mathbf{b}} \cdot \nabla B \frac{\partial h_s}{\partial \xi'}. \quad (\text{B.85})$$

# Appendix C

## Differential geometry

### C.1 Cartesian results

Physical space can be labelled with Cartesian axes, with coordinates  $x, y$  and  $z$ . The basis vectors for these coordinates are  $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ , which are orthonormal. A point in space  $\mathbf{r}$  can therefore be labelled

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}. \quad (\text{C.1})$$

The differential of this position is

$$\begin{aligned} d\mathbf{r} &= dx \frac{\partial \mathbf{r}}{\partial x} + dy \frac{\partial \mathbf{r}}{\partial y} + dz \frac{\partial \mathbf{r}}{\partial z} \\ &= dx\hat{\mathbf{i}} + dy\hat{\mathbf{j}} + dz\hat{\mathbf{k}} \end{aligned} \quad (\text{C.2})$$

and so we see the Cartesian basis vectors can be written via  $\hat{\mathbf{i}} = \partial \mathbf{r} / \partial x$ ,  $\hat{\mathbf{j}} = \partial \mathbf{r} / \partial y$  and  $\hat{\mathbf{k}} = \partial \mathbf{r} / \partial z$ .

A general vector  $\mathbf{V}$  can be written

$$\mathbf{V} = V_x \hat{\mathbf{i}} + V_y \hat{\mathbf{j}} + V_z \hat{\mathbf{k}} \quad (\text{C.3})$$

where  $V_x = \mathbf{V} \cdot \hat{\mathbf{i}}$ ,  $V_y = \mathbf{V} \cdot \hat{\mathbf{j}}$  and  $V_z = \mathbf{V} \cdot \hat{\mathbf{k}}$  are the Cartesian components of  $\mathbf{V}$ . The scalar product between two vectors  $\mathbf{A}$  and  $\mathbf{B}$  is

$$\mathbf{A} \cdot \mathbf{B} = A_x B_x + A_y B_y + A_z B_z \quad (\text{C.4})$$

and the vector product is

$$\mathbf{A} \times \mathbf{B} = (A_y B_z - A_z B_y) \hat{\mathbf{i}} + (A_z B_x - A_x B_z) \hat{\mathbf{j}} + (A_x B_y - A_y B_x) \hat{\mathbf{k}}. \quad (\text{C.5})$$

For a spatially-dependent function  $f = f(x, y, z)$ , its derivative with respect to another function  $g$  is obtained via the chain rule,

$$\frac{\partial f}{\partial g} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial g} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial g} + \frac{\partial f}{\partial z} \frac{\partial z}{\partial g} \quad (\text{C.6})$$

We now consider the directional derivatives of  $f$ . Defining the gradient of a function, we have

$$\nabla f = \frac{\partial f}{\partial x} \hat{\mathbf{i}} + \frac{\partial f}{\partial y} \hat{\mathbf{j}} + \frac{\partial f}{\partial z} \hat{\mathbf{k}} \quad (\text{C.7})$$

Note that one may write  $\hat{\mathbf{i}} = \nabla x$ ,  $\hat{\mathbf{j}} = \nabla y$  and  $\hat{\mathbf{k}} = \nabla z$ .

The divergence of a vector field is defined by

$$\nabla \cdot \mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} \quad (\text{C.8})$$

and its curl is

$$\nabla \times \mathbf{V} = \left( \frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z} \right) \hat{\mathbf{i}} + \left( \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x} \right) \hat{\mathbf{j}} + \left( \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \right) \hat{\mathbf{k}}. \quad (\text{C.9})$$

We now show some important results involving  $\nabla$ . The first is that the curl of a gradient is trivially zero, by using  $\mathbf{V} = \nabla f$  in equation C.9:

$$\begin{aligned} \nabla \times \nabla f &= \left( \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial z} \right) - \frac{\partial}{\partial z} \left( \frac{\partial f}{\partial y} \right) \right) \hat{\mathbf{i}} + \left( \frac{\partial}{\partial z} \left( \frac{\partial f}{\partial x} \right) - \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial z} \right) \right) \hat{\mathbf{j}} \\ &\quad + \left( \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) - \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) \right) \hat{\mathbf{k}} \\ &= \mathbf{0} \end{aligned} \quad (\text{C.10})$$

by the commutativity of the partial derivatives. Therefore any vector which has zero curl can be expressed as the gradient of a scalar function

$$\nabla \times \mathbf{V} = \mathbf{0} \implies \mathbf{V} = \nabla f. \quad (\text{C.11})$$

A second result comes from combining equations C.8 and C.9, to show that the divergence of a curl is also trivially zero:

$$\begin{aligned} \nabla \cdot (\nabla \times \mathbf{V}) &= \frac{\partial}{\partial x} \left( \frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z} \right) + \frac{\partial}{\partial y} \left( \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x} \right) + \frac{\partial}{\partial z} \left( \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \right) \\ &= 0 \end{aligned} \quad (\text{C.12})$$

again by the commutativity of the partial derivatives. Thus a divergenceless vector can be expressed as a curl of a vector function

$$\nabla \cdot \mathbf{V} = 0 \implies \mathbf{V} = \nabla \times \mathbf{A}. \quad (\text{C.13})$$

These relations can then be used to show a third result, namely that the divergence of the cross product of two gradients is also trivially zero. Because equation C.12 applies to any vector  $\mathbf{V}$ , consider a vector of the form  $\mathbf{V} = F\nabla G$ , for scalar functions  $F$  and  $G$ . Using the chain rule, the curl of  $F\nabla G$  is

$$\nabla \times (F\nabla G) = \nabla F \times \nabla G + F(\nabla \times \nabla G) \quad \mathbf{0} \quad (\text{C.14})$$

where the second term is trivially zero by equation C.10. Inserting C.14 into equation C.12, one obtains

$$\nabla \cdot (\nabla F \times \nabla G) = 0. \quad (\text{C.15})$$

Therefore any divergenceless vector can either be expressed as the curl of a vector, or equivalently as the cross product of two gradients

$$\begin{aligned} \nabla \cdot \mathbf{V} = 0 &\implies \mathbf{V} = \nabla \times \mathbf{A} \\ \mathbf{V} &= \nabla F \times \nabla G. \end{aligned} \quad (\text{C.16})$$

Note that because  $\nabla \times \nabla f = \mathbf{0}$ ,  $\nabla \cdot (\nabla \times \mathbf{V}) = \mathbf{0}$  and  $\nabla \cdot (\nabla F \times \nabla G) = 0$  are written without explicit reference to coordinates, they hold in any coordinate system.

## C.2 General coordinate systems

### C.2.1 Basis vectors

#### Covariant basis vectors

Transforming from Cartesian coordinates to a new set of coordinates, we label these  $\xi^i$  with  $i = \{1, 2, 3\}$ , where a superscript is used for coordinates by convention. We then have for a position vector  $\mathbf{r} = \mathbf{r}(x(\xi^i), y(\xi^i), z(\xi^i)) = \mathbf{r}(\xi^1, \xi^2, \xi^3)$ , such that

$$\mathbf{r} = x(\xi^i) \hat{\mathbf{i}} + y(\xi^i) \hat{\mathbf{j}} + z(\xi^i) \hat{\mathbf{k}}. \quad (\text{C.17})$$

Its differential in the new coordinates is

$$\begin{aligned} d\mathbf{r} &= \frac{\partial \mathbf{r}}{\partial \xi^1} d\xi^1 + \frac{\partial \mathbf{r}}{\partial \xi^2} d\xi^2 + \frac{\partial \mathbf{r}}{\partial \xi^3} d\xi^3 \\ &= \sum_{i=1}^3 \frac{\partial \mathbf{r}}{\partial \xi^i} d\xi^i. \end{aligned} \quad (\text{C.18})$$

To write this more compactly we use the Einstein Summation Convention. This omits summation signs but implies summing over an index when it is repeated, which should only occur when one index is ‘up’ and one is ‘down’. A given index may therefore only appear a maximum of twice. Equation C.18 can hence be written

$$d\mathbf{r} = \frac{\partial \mathbf{r}}{\partial \xi^i} d\xi^i. \quad (\text{C.19})$$

Note that the label for an index that is summed over is arbitrary, and so any symbol can replace  $i$  in equation C.19 without changing its meaning.

Equation C.19 introduces three ‘natural’ vectors for the system,  $\partial\mathbf{r}/\partial\xi^i$ , which are called the *covariant* basis vectors. In this work the shorthand  $\mathbf{a}_i = \partial\mathbf{r}/\partial\xi^i$  is used. Using the chain rule, these are related to the Cartesian system via

$$\begin{aligned}\mathbf{a}_i &= \frac{\partial\mathbf{r}}{\partial\xi^i} = \frac{\partial x}{\partial\xi^i} \frac{\partial\mathbf{r}}{\partial x} + \frac{\partial y}{\partial\xi^i} \frac{\partial\mathbf{r}}{\partial y} + \frac{\partial z}{\partial\xi^i} \frac{\partial\mathbf{r}}{\partial z} \\ &= \frac{\partial x}{\partial\xi^i} \hat{\mathbf{i}} + \frac{\partial y}{\partial\xi^i} \hat{\mathbf{j}} + \frac{\partial z}{\partial\xi^i} \hat{\mathbf{k}}.\end{aligned}\tag{C.20}$$

Note that unlike the basis vectors of Cartesian coordinates, the covariant basis vectors of a given system are not necessarily unit vectors, nor are they necessarily orthogonal to one another.

### Contravariant basis vectors

Now consider a spatially-dependent function  $f = f(\xi^1, \xi^2, \xi^3)$ . Its partial derivative with respect to some other function  $g$  in terms of the new basis is

$$\frac{\partial f}{\partial g} = \frac{\partial f}{\partial\xi^i} \frac{\partial\xi^i}{\partial g}\tag{C.21}$$

which is simply the chain rule in the new system.

Now consider the directional derivatives of  $f$ . Starting from the definition of the gradient, shown in equation C.7, it follows that

$$\begin{aligned}\nabla f &= \frac{\partial f}{\partial x} \hat{\mathbf{i}} + \frac{\partial f}{\partial y} \hat{\mathbf{j}} + \frac{\partial f}{\partial z} \hat{\mathbf{k}} \\ &= \frac{\partial f}{\partial\xi^i} \frac{\partial\xi^i}{\partial x} \hat{\mathbf{i}} + \frac{\partial f}{\partial\xi^i} \frac{\partial\xi^i}{\partial y} \hat{\mathbf{j}} + \frac{\partial f}{\partial\xi^i} \frac{\partial\xi^i}{\partial z} \hat{\mathbf{k}} \\ &= \frac{\partial f}{\partial\xi^i} \nabla\xi^i\end{aligned}\tag{C.22}$$

where we have used equation C.21, and the final line is the gradient of a scalar expressed in our new coordinate system.

Equation C.22 introduces another set of ‘natural’ vectors,  $\nabla\xi^i$ , which are named the *contravariant* basis vectors. These are related to the Cartesian system via

$$\nabla\xi^i = \frac{\partial\xi^i}{\partial x} \hat{\mathbf{i}} + \frac{\partial\xi^i}{\partial y} \hat{\mathbf{j}} + \frac{\partial\xi^i}{\partial z} \hat{\mathbf{k}}.\tag{C.23}$$

Like the covariant basis vectors, these are not necessarily unit vectors or orthogonal.

### Basis vector relations and the Jacobian

We now consider how the basis vectors relate to one another. Consider the scalar product between a covariant and a contravariant basis vector,  $\mathbf{a}_i \cdot \nabla\xi^j$ . Using the chain rule to express

them in Cartesian coordinates,

$$\begin{aligned}
\mathbf{a}_i \cdot \nabla \xi^j &= \left( \frac{\partial x}{\partial \xi^i} \hat{\mathbf{i}} + \frac{\partial y}{\partial \xi^i} \hat{\mathbf{j}} + \frac{\partial z}{\partial \xi^i} \hat{\mathbf{k}} \right) \cdot \left( \frac{\partial \xi^j}{\partial x} \hat{\mathbf{i}} + \frac{\partial \xi^j}{\partial y} \hat{\mathbf{j}} + \frac{\partial \xi^j}{\partial z} \hat{\mathbf{k}} \right) \\
&= \frac{\partial \xi^j}{\partial x} \frac{\partial x}{\partial \xi^i} + \frac{\partial \xi^j}{\partial y} \frac{\partial y}{\partial \xi^i} + \frac{\partial \xi^j}{\partial z} \frac{\partial z}{\partial \xi^i} \\
&= \frac{\partial \xi^j}{\partial \xi^i} \\
&= \delta_i^j
\end{aligned} \tag{C.24}$$

where we have used equation C.6, and  $\delta_i^j$  is the Kronecker delta, which is equal to 1 for  $i = j$ , and 0 for  $i \neq j$ . This condition allows the basis vectors of one type to be written in terms of those of the other.

Consider  $\nabla \xi^1$ . Because we know this is orthogonal to  $\mathbf{a}_2$  and  $\mathbf{a}_3$ , we may write  $\nabla \xi^1 = C \mathbf{a}_2 \times \mathbf{a}_3$ , with proportionality factor  $C$ . By taking the scalar product of  $\mathbf{a}_1$  with both sides, we find  $C = 1 / [\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)]$ , and thus we also may write  $\nabla \xi^2 = C \mathbf{a}_3 \times \mathbf{a}_1$  and  $\nabla \xi^3 = C \mathbf{a}_1 \times \mathbf{a}_2$ . Similarly, we can write  $\mathbf{a}_1 = D \nabla \xi^2 \times \nabla \xi^3$ , from which we find  $D = 1 / [\nabla \xi^1 \cdot (\nabla \xi^2 \times \nabla \xi^3)]$  and thus  $\mathbf{a}_2 = D \nabla \xi^3 \times \nabla \xi^1$  and  $\mathbf{a}_3 = D \nabla \xi^1 \times \nabla \xi^2$ . To relate  $C$  and  $D$  to one another, we write

$$\begin{aligned}
\nabla \xi^1 &= C \mathbf{a}_2 \times \mathbf{a}_3 \\
&= CD^2 [(\nabla \xi^3 \times \nabla \xi^1) \times (\nabla \xi^1 \times \nabla \xi^2)] \\
&= CD^2 \left[ \nabla \xi^1 (\nabla \xi^2 \cdot (\nabla \xi^3 \times \nabla \xi^1)) - \nabla \xi^2 (\nabla \xi^1 \cdot (\nabla \xi^3 \times \nabla \xi^1)) \right] \mathbf{10} \\
&= CD \nabla \xi^1
\end{aligned} \tag{C.25}$$

where we have used the triple vector product. Hence  $C = 1/D$ , and so

$$\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3) = \frac{1}{\nabla \xi^1 \cdot (\nabla \xi^2 \times \nabla \xi^3)} = \mathcal{J} \tag{C.26}$$

where we have defined the Jacobian for the system,  $\mathcal{J}$ . Making use of the Levi-Civita symbol<sup>1</sup>, the basis vectors may therefore be written

$$\nabla \xi^i = \frac{1}{2\mathcal{J}} \epsilon^{ijk} \mathbf{a}_j \times \mathbf{a}_k \tag{C.27}$$

$$\mathbf{a}_i = \frac{1}{2} \mathcal{J} \epsilon_{ijk} \nabla \xi^j \times \nabla \xi^k \tag{C.28}$$

---

<sup>1</sup>This mathematical object can be written with raised indices  $\epsilon^{ijk}$  or lowered indices  $\epsilon_{ijk}$ , and is defined as equal to 1 for even permutations of the indices, -1 for odd permutations, and 0 if any indices are repeated.

where the factor of 1/2 is necessary to nullify double-counting from the Levi-Civita symbol. For example, evaluating equation C.27 for  $i = 1$ , we get

$$\begin{aligned}\nabla\xi^1 &= \frac{1}{2\mathcal{J}}(\mathbf{a}_2 \times \mathbf{a}_3 - \mathbf{a}_3 \times \mathbf{a}_2) \\ &= \frac{1}{\mathcal{J}}(\mathbf{a}_2 \times \mathbf{a}_3).\end{aligned}\tag{C.29}$$

We may also write

$$\nabla\xi^j \times \nabla\xi^k = \frac{1}{\mathcal{J}}\epsilon^{ijk}\mathbf{a}_i\tag{C.30}$$

$$\mathbf{a}_j \times \mathbf{a}_k = \mathcal{J}\epsilon_{ijk}\nabla\xi^i.\tag{C.31}$$

Note there is *not* double counting in this case, as once the two left-hand side indices have been specified, there is only one non-zero value of the Levi-Civita symbol remaining.

## C.2.2 Vector representations

### Vector components and the metric

The existence of two types of basis vector for a general system, the contravariant basis and the covariant basis, allows any general vector to be expressed in *either* basis,

$$\mathbf{V} = V_i\nabla\xi^i = V^j\mathbf{a}_j.\tag{C.32}$$

where we have defined the covariant components of  $\mathbf{V}$ ,  $V_i = \mathbf{V} \cdot \mathbf{a}_i$ , and the contravariant components,  $V^i = \mathbf{V} \cdot \nabla\xi^i$ . This equivalence allows one to express one type of vector components in terms of the other.

Consider taking the dot product of  $\mathbf{V}$  with  $\mathbf{a}_k$ . The covariant component expression simply becomes  $\mathbf{V} \cdot \mathbf{a}_k = V_i\nabla\xi^i \cdot \mathbf{a}_k = V_i\delta_k^i = V_k$ , as expected. Applied to the contravariant component expression, one has  $\mathbf{V} \cdot \mathbf{a}_k = V^j(\mathbf{a}_j \cdot \mathbf{a}_k) = V^jg_{jk}$ . Here we introduce the *metric*, defined such that  $g_{ij} = \mathbf{a}_i \cdot \mathbf{a}_j$ . Given its lowered indices this can be referred to as the covariant metric. The metric is even with respect to its indices,  $g_{ij} = g_{ji}$ , by the symmetry of the dot product. We therefore find that the covariant components of a vector can be expressed as a sum over the contravariant components, multiplied by the appropriate metric components.

$$V_i = g_{ij}V^j.\tag{C.33}$$

Inserting this result into the covariant component expression of equation C.32 one finds

$$\begin{aligned}\mathbf{V} &= V_i\nabla\xi^i \\ &= g_{ij}V^j\nabla\xi^i \\ &= V^j(g_{ij}\nabla\xi^i)\end{aligned}\tag{C.34}$$

which, via comparison with the contravariant expression, implies that

$$\mathbf{a}_i = g_{ij} \nabla \xi^j. \quad (\text{C.35})$$

and so the covariant basis vectors can be expressed as a sum over the contravariant basis vectors multiplying terms in the metric.

Similarly, we can define an object  $g^{ij} = \nabla \xi^i \cdot \nabla \xi^j$ , dubbed the contravariant metric, from which we find using analogous analysis that

$$V^i = g^{ij} V_j \quad (\text{C.36})$$

$$\nabla \xi^i = g^{ij} \mathbf{a}_j. \quad (\text{C.37})$$

We see that one can therefore change the type of basis or coordinate used by ‘raising/lowering’ the indices via the metric.

One can consider the metric as a matrix object. Writing equations C.33 and C.36 as a system of linear equations, one has

$$V_i = g_{ij} V^j \rightarrow \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix} \begin{pmatrix} V^1 \\ V^2 \\ V^3 \end{pmatrix} \quad (\text{C.38})$$

$$V^i = g^{ij} V_j \rightarrow \begin{pmatrix} V^1 \\ V^2 \\ V^3 \end{pmatrix} = \begin{pmatrix} g^{11} & g^{12} & g^{13} \\ g^{21} & g^{22} & g^{23} \\ g^{31} & g^{32} & g^{33} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} \quad (\text{C.39})$$

from which we see that the contravariant metric  $g^{ij}$  is the inverse of the covariant metric,  $g_{ij}$ , and thus can be referred to as the inverse metric. This can also be seen from

$$V_i = g_{ij} V^j = g_{ij} g^{jk} V_k \quad (\text{C.40})$$

from which one must conclude that

$$g_{ij} g^{jk} = \delta_i^k \quad (\text{C.41})$$

which is the matrix equation for inverses.

## Metric results

We shall now show how the determinant of the metric is related to the Jacobian. Using equation C.31,

$$\begin{aligned} \mathcal{J} \epsilon_{ijk} &= \mathbf{a}_i \cdot (\mathbf{a}_j \times \mathbf{a}_k) \\ &= g_{il} g_{jm} g_{kn} \left[ \nabla \xi^l \cdot (\nabla \xi^m \times \nabla \xi^n) \right] \\ &= g_{il} g_{jm} g_{kn} \epsilon^{lmn} \frac{1}{\mathcal{J}}. \end{aligned} \quad (\text{C.42})$$

Multiplying both sides by the Jacobian, and taking the coordinates  $i = 1, j = 2, k = 3$ , one finds

$$\mathcal{J}^2 = \epsilon^{lmn} g_{1l} g_{2m} g_{3n} \quad (\text{C.43})$$

which, considering the metric as a matrix, is the equation for the determinant. Hence we find

$$\det [g_{ij}] = \mathcal{J}^2. \quad (\text{C.44})$$

Either by the rule of the determinant of inverse matrices, or from similar analysis using equation C.30, one finds

$$\det [g^{ij}] = \frac{1}{\mathcal{J}^2}. \quad (\text{C.45})$$

We will now relate the components of the metric and the inverse metric to one another. From the definition of the components of the contravariant metric, we have

$$\begin{aligned} g^{ij} &= \nabla \xi^i \cdot \nabla \xi^j \\ &= \left( \frac{1}{2\mathcal{J}} \right)^2 \epsilon^{imn} \epsilon^{j pq} (\mathbf{a}_m \times \mathbf{a}_n) \cdot (\mathbf{a}_p \times \mathbf{a}_q) \\ &= \left( \frac{1}{2\mathcal{J}} \right)^2 \epsilon^{imn} \epsilon^{j pq} (g_{mp} g_{nq} - g_{mq} g_{np}) \end{aligned} \quad (\text{C.46})$$

where we have used equation C.27 and an appropriate vector identity<sup>2</sup>. Consider terms involved in the final summation,  $-\epsilon^{j pq} g_{mq} g_{np}$ . Because  $p$  and  $q$  are arbitrary labels due to them being summed over, let us relabel them to  $r$  and  $s$ ,  $-\epsilon^{j rs} g_{ms} g_{nr}$ . Now we swap the  $r$  and  $s$  indices in the Levi-Civita symbol, using  $\epsilon^{j rs} = -\epsilon^{j sr}$ , such that we have  $-\epsilon^{j rs} g_{ms} g_{nr} = \epsilon^{j sr} g_{ms} g_{nr}$ . Again using the freedom of the arbitrary index labels, we then relabel  $s$  as  $p$  and  $r$  as  $q$  on the right-hand side, but  $r$  as  $p$  and  $s$  as  $q$  on the left-hand side, giving  $-\epsilon^{j pq} g_{mq} g_{np} = \epsilon^{j pq} g_{mp} g_{nq}$ . Using this in equation C.46 yields

$$g^{ij} = \frac{1}{2\mathcal{J}^2} \epsilon^{imn} \epsilon^{j pq} g_{mp} g_{nq} \quad (\text{C.47})$$

which is our result. Note again that the factor of  $1/2$  prevents double counting. For example, calculating  $g^{11}$ , one has

$$\begin{aligned} g^{11} &= \frac{1}{2\mathcal{J}^2} [(\epsilon^{123} \epsilon^{123} + \epsilon^{132} \epsilon^{132}) g_{22} g_{33} + \epsilon^{132} \epsilon^{123} g_{32} g_{23} + \epsilon^{123} \epsilon^{132} g_{23} g_{32}] \\ &= \frac{1}{2\mathcal{J}^2} [2g_{22} g_{33} - 2(g_{23})^2] \\ &= \frac{g_{22} g_{33} - (g_{23})^2}{\mathcal{J}^2} \end{aligned} \quad (\text{C.48})$$

---

<sup>2</sup>It can be shown that  $(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{B} \cdot \mathbf{C})(\mathbf{A} \cdot \mathbf{D})$ .

where we have used the symmetry of the metric,  $g_{23} = g_{32}$ .

Analogous analysis using equation C.28 finds

$$g_{ij} = \frac{\mathcal{J}^2}{2} \epsilon_{imn} \epsilon_{jpq} g^{mp} g^{nq}. \quad (\text{C.49})$$

Evaluating the contravariant metric terms explicitly, one has

$$g^{ij} = \frac{1}{\mathcal{J}^2} \begin{pmatrix} g_{22}g_{33} - (g_{23})^2 & g_{23}g_{13} - g_{12}g_{33} & g_{12}g_{23} - g_{13}g_{22} \\ g_{23}g_{13} - g_{12}g_{33} & g_{11}g_{33} - (g_{13})^2 & g_{12}g_{13} - g_{11}g_{23} \\ g_{12}g_{23} - g_{13}g_{22} & g_{12}g_{13} - g_{11}g_{23} & g_{11}g_{22} - (g_{12})^2 \end{pmatrix} \quad (\text{C.50})$$

and correspondingly for the covariant metric

$$g_{ij} = \mathcal{J}^2 \begin{pmatrix} g^{22}g^{33} - (g^{23})^2 & g^{23}g^{13} - g^{12}g^{33} & g^{12}g^{23} - g^{13}g^{22} \\ g^{23}g^{13} - g^{12}g^{33} & g^{11}g^{33} - (g^{13})^2 & g^{12}g^{13} - g^{11}g^{23} \\ g^{12}g^{23} - g^{13}g^{22} & g^{12}g^{13} - g^{11}g^{23} & g^{11}g^{22} - (g^{12})^2 \end{pmatrix}. \quad (\text{C.51})$$

### C.2.3 Vector operations

We have seen from equation C.32 that for a general coordinate system, vectors can be represented in two ways. The decision on which representation to use can be influenced by the problem at hand.

#### Vector algebra

First we consider the scalar product of  $\mathbf{V}$  and  $\mathbf{W}$ , given for example by

$$\mathbf{V} \cdot \mathbf{W} = (V^i \mathbf{a}_i) \cdot (W_j \nabla \xi^j) = V^i W_i. \quad (\text{C.52})$$

We note that this can equivalently be expressed via  $V^i W_i = g^{ij} V_j W_i = V_j W^j = g_{ij} V^i W^j$ .

For the vector product, defining  $\mathbf{U} = \mathbf{V} \times \mathbf{W}$ , one has

$$\mathbf{U} = V_j W_k \nabla \xi^j \times \nabla \xi^k = \frac{1}{\mathcal{J}} \epsilon^{ijk} V_j W_k \mathbf{a}_i \quad (\text{C.53})$$

and thus the contravariant components of  $\mathbf{U}$  are

$$U^i = \frac{1}{\mathcal{J}} \epsilon^{ijk} V_j W_k. \quad (\text{C.54})$$

Alternatively, in the corresponding representation,

$$\mathbf{U} = V^j W^k \mathbf{a}_j \times \mathbf{a}_k = \mathcal{J} \epsilon_{ijk} V^j W^k \nabla \xi^i \quad (\text{C.55})$$

and thus the covariant components of  $\mathbf{U}$  are

$$U_i = \mathcal{J} \epsilon_{ijk} V^j W^k. \quad (\text{C.56})$$

### Vector calculus

We now consider operations involving  $\nabla$ . If a vector can be expressed as a gradient of a scalar,  $\mathbf{V} = \nabla f$ , we have

$$\mathbf{V} = \frac{\partial f}{\partial \xi^i} \nabla \xi^i \quad (\text{C.57})$$

and so the covariant components are more ‘natural’, with  $V_i = \partial f / \partial \xi^i$ . The contravariant components are  $V^i = g^{ij} \partial f / \partial \xi^j$ .

Next we calculate the curl of a vector. Using the covariant component representation, one finds

$$\begin{aligned} \nabla \times \mathbf{V} &= \nabla \times (V_k \nabla \xi^k) \\ &= \nabla V_k \times \nabla \xi^k + V_k \left( \nabla \times \nabla \xi^k \right) \mathbf{0} \\ &= \frac{\partial V_k}{\partial \xi^j} \nabla \xi^j \times \nabla \xi^k \\ &= \frac{1}{\mathcal{J}} \epsilon^{ijk} \frac{\partial V_k}{\partial \xi^j} \mathbf{a}_i \end{aligned} \quad (\text{C.58})$$

where we have used equation C.10.

Finally for the divergence, using the contravariant components, one has

$$\begin{aligned} \nabla \cdot \mathbf{V} &= \nabla \cdot (V^i \mathbf{a}_i) \\ &= \nabla \cdot \left( V^i \frac{1}{2} \mathcal{J} \epsilon_{ijk} \nabla \xi^j \times \nabla \xi^k \right) \\ &= \frac{1}{2} \epsilon_{ijk} [\nabla (V^i \mathcal{J})] \cdot (\nabla \xi^j \times \nabla \xi^k) + \frac{1}{2} \epsilon_{ijk} V^i \mathcal{J} \nabla \cdot \left( \nabla \xi^j \times \nabla \xi^k \right) \mathbf{0} \\ &= \frac{1}{\mathcal{J}} [\nabla (V^i \mathcal{J})] \cdot \mathbf{a}_i \\ &= \frac{1}{\mathcal{J}} \frac{\partial (\mathcal{J} V^i)}{\partial \xi^i} \end{aligned} \quad (\text{C.59})$$

where we have used equations C.15 and C.28.

Subsequent vector operations can be obtained from combinations of the above. For example the Laplacian, using  $\mathbf{V} = \nabla f$  in equation C.59, is

$$\begin{aligned} \nabla^2 f &= \nabla \cdot \nabla f \\ &= \frac{1}{\mathcal{J}} \frac{\partial}{\partial \xi^i} (\mathcal{J} [\nabla f \cdot \nabla \xi^i]) \\ &= \frac{1}{\mathcal{J}} \frac{\partial}{\partial \xi^i} \left( \mathcal{J} g^{ij} \frac{\partial f}{\partial \xi^j} \right). \end{aligned} \quad (\text{C.60})$$

### C.2.4 Transforming between systems

Let us suppose we are using a coordinate system  $\{\xi^i\}$ , and transform to a new system  $\{\eta^j\}$ . Priming the  $\xi^i$  system for clarity,  $\xi^i \rightarrow \xi'^i$ , the contravariant basis vectors in the new system are

$$\nabla\eta^j = \frac{\partial\eta^j}{\partial\xi'^i} \nabla\xi'^i \quad (\text{C.61})$$

from equation C.22, and the covariant basis vectors are

$$\begin{aligned} \mathbf{a}_j &= \frac{\partial\mathbf{r}}{\partial\eta^j} \\ &= \frac{\partial\mathbf{r}}{\partial\xi'^i} \frac{\partial\xi'^i}{\partial\eta^j} \\ &= \frac{\partial\xi'^i}{\partial\eta^j} \mathbf{a}'_i. \end{aligned} \quad (\text{C.62})$$

Using these results, the components of a general vector  $\mathbf{V}$  are

$$V^j = \mathbf{V} \cdot \nabla\eta^j = \frac{\partial\eta^j}{\partial\xi'^i} V'^i \quad (\text{C.63})$$

$$V_j = \mathbf{V} \cdot \mathbf{a}_j = \frac{\partial\xi'^i}{\partial\eta^j} V'_i \quad (\text{C.64})$$

with metric components

$$g_{ij} = \frac{\partial\xi'^m}{\partial\eta^i} \frac{\partial\xi'^m}{\partial\eta^j} g'_{mn}, \quad g^{ij} = \frac{\partial\eta^i}{\partial\xi'^m} \frac{\partial\eta^j}{\partial\xi'^m} g'^{mn}. \quad (\text{C.65})$$

# Appendix D

## Equilibrium geometry

### D.1 Equilibrium plasma momentum equation

Taking the fluid equation for the plasma momentum, equation 1.25, and summing over species  $s$  we obtain

$$\sum_s m_s n_s \frac{\partial \mathbf{u}_s}{\partial t} + \sum_s m_s n_s (\mathbf{u}_s \cdot \nabla) \mathbf{u}_s + \nabla P + \nabla \cdot \underline{\underline{\pi}} - \rho \mathbf{E} - \mathbf{J} \times \mathbf{B} = \mathbf{0} \quad (\text{D.1})$$

where we have defined the total plasma pressure  $P = \sum_s P_s$ , the total anisotropic pressure tensor  $\underline{\underline{\pi}} = \sum_s \underline{\underline{\pi}}_s$ , and used the definitions of the charge density  $\rho = \sum_s q_s n_s$  and the current density  $\mathbf{J} = \sum_s q_s n_s \mathbf{u}_s$ . The collision terms exactly cancel when summed over species, as there can be no net momentum change arising from inter-species collisions [9],

$$\sum_s \int m_s \mathbf{v} C_s d^3 \mathbf{v} = \mathbf{0}. \quad (\text{D.2})$$

Ensemble averaging equation D.1 in the limit of small plasma flows, we have to lowest order

$$\mathbf{J} \times \mathbf{B} = \nabla P \quad (\text{D.3})$$

where because the lowest order distribution function of each species is a Maxwellian from equation 3.92, the equilibrium pressure is isotropic  $\underline{\underline{\pi}}_s = \underline{\underline{0}} \implies \underline{\underline{\pi}} = \underline{\underline{0}}$ , and the electric field term is zero due to lowest order quasineutrality, equation 3.99. Note that the ordering assumptions used imply that the size of the equilibrium current density is of order  $|\mathbf{J}| \sim \mathcal{O}(n_0 T_0 / (aB)) \sim \mathcal{O}(\rho_* e n_0 v_{\text{th}})$ .

## D.2 Cylindrical geometry

The transformations between Cartesian coordinates and a cylindrical system  $\{R, \Phi, Z\}$  are

$$R = \sqrt{x^2 + y^2} \qquad x = R \cos(\Phi) \qquad (\text{D.4})$$

$$\Phi = \arctan\left(\frac{y}{x}\right) \qquad y = R \sin(\Phi) \qquad (\text{D.5})$$

$$Z = z \qquad z = Z \qquad (\text{D.6})$$

with contravariant and covariant basis vectors

$$\nabla R = \cos(\Phi) \hat{\mathbf{i}} + \sin(\Phi) \hat{\mathbf{j}} \qquad \mathbf{a}_R = \cos(\Phi) \hat{\mathbf{i}} + \sin(\Phi) \hat{\mathbf{j}} \qquad (\text{D.7})$$

$$\nabla \Phi = \frac{1}{R} \left( -\sin(\Phi) \hat{\mathbf{i}} + \cos(\Phi) \hat{\mathbf{j}} \right) \qquad \mathbf{a}_\Phi = R \left( -\sin(\Phi) \hat{\mathbf{i}} + \cos(\Phi) \hat{\mathbf{j}} \right) \qquad (\text{D.8})$$

$$\nabla Z = \hat{\mathbf{k}} \qquad \mathbf{a}_Z = \hat{\mathbf{k}} \qquad (\text{D.9})$$

and metric components

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & R^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad g^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/R^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \qquad (\text{D.10})$$

The Jacobian for the system is  $\mathcal{J}_{R\Phi Z} = R$ .

### D.2.1 Axisymmetric divergencelessness

Using equation C.59, the divergence of the magnetic field in the cylindrical system is

$$\frac{1}{R} \frac{\partial (RB^R)}{\partial R} + \frac{1}{R} \frac{\partial (RB^\Phi)}{\partial \Phi} + \frac{1}{R} \frac{\partial (RB^Z)}{\partial Z} = 0 \qquad (\text{D.11})$$

where the second term is zero due to axisymmetry. This implies that the divergence of just the poloidal part of the field is zero, and so can be written as the cross product of two gradients from equation C.15,

$$\nabla \cdot (B^R \mathbf{a}_R + B^Z \mathbf{a}_Z) = 0 \implies B^R \mathbf{a}_R + B^Z \mathbf{a}_Z = \nabla f_p \times \nabla g_p \qquad (\text{D.12})$$

for two potentials  $f_p$  and  $g_p$ . For full generality we may express one of these potentials as a univariate function of another function, which we call  $\varrho$ , such that  $f_p = f_p(\varrho)$ . This gives

$$B^R \mathbf{a}_R + B^Z \mathbf{a}_Z = \frac{df_p}{d\varrho} \nabla \varrho \times \nabla g_p. \qquad (\text{D.13})$$

To solve for these new potentials, we write the covariant basis vectors in equation D.13 in terms of the contravariant basis vectors using equation C.28,  $\mathbf{a}_R = -R\nabla Z \times \nabla\Phi$  and  $\mathbf{a}_Z = R\nabla R \times \nabla\Phi$ :

$$[RB^Z\nabla R - RB^R\nabla Z] \times \nabla\Phi = \frac{df_p}{d\varrho}\nabla\varrho \times \nabla g_p \quad (\text{D.14})$$

from which we obtain  $g_p = \Phi$ , and

$$\frac{df_p}{d\varrho}\nabla\varrho = RB^Z\nabla R - RB^R\nabla Z. \quad (\text{D.15})$$

The partial derivatives of  $\varrho$  therefore satisfy

$$\frac{df_p}{d\varrho}\frac{\partial\varrho}{\partial R} = RB^Z \quad (\text{D.16})$$

$$\frac{df_p}{d\varrho}\frac{\partial\varrho}{\partial\Phi} = 0 \quad (\text{D.17})$$

$$\frac{df_p}{d\varrho}\frac{\partial\varrho}{\partial Z} = -RB^R \quad (\text{D.18})$$

and so we have found that  $\varrho$  is independent of toroidal angle,  $\varrho = \varrho(R, Z)$ . The total equilibrium magnetic field in the cylindrical system can now be written

$$\mathbf{B} = \frac{df_p}{d\varrho}\nabla\varrho \times \nabla\Phi + B^\Phi\mathbf{a}_\Phi \quad (\text{D.19})$$

which, upon taking the scalar product with  $\nabla\varrho$  gives  $\mathbf{B} \cdot \nabla\varrho = 0$ , implying that  $\varrho$  can be used as one of the potentials in equation 3.108. Choosing  $F = \varrho$ ,

$$\mathbf{B} = \nabla\varrho \times \nabla G. \quad (\text{D.20})$$

Upon evaluating the gradient of  $G$  in the cylindrical system, one obtains

$$\mathbf{B} = \frac{\partial G}{\partial\Phi}\nabla\varrho \times \nabla\Phi + \nabla\varrho \times \left( \frac{\partial G}{\partial R}\nabla R + \frac{\partial G}{\partial Z}\nabla Z \right) \quad (\text{D.21})$$

which, via comparison with equation D.19, implies  $\partial G/\partial\Phi = df_p/d\varrho$ , which we relabel to  $\bar{c}_1(\varrho)$ . The potential  $G$  therefore has the solution

$$G(R, \Phi, Z) = \bar{c}_1(\varrho)\Phi + \tilde{G}(R, Z). \quad (\text{D.22})$$

### D.3 Toroidal coordinates

We now transform from cylindrical coordinates to toroidal coordinates, defined by

$$\varrho = \varrho(R, Z) \quad R = R(\varrho, \vartheta) \quad (\text{D.23})$$

$$\vartheta = \vartheta(R, Z) \quad \Phi = -\zeta \quad (\text{D.24})$$

$$\zeta = -\Phi \quad Z = Z(\varrho, \vartheta). \quad (\text{D.25})$$

The basis vectors are

$$\nabla_{\varrho} = \frac{\partial \varrho}{\partial R} \nabla R + \frac{\partial \varrho}{\partial Z} \nabla Z \quad \mathbf{a}_{\varrho} = \frac{\partial R}{\partial \varrho} \mathbf{a}_R + \frac{\partial Z}{\partial \varrho} \mathbf{a}_Z \quad (\text{D.26})$$

$$\nabla_{\vartheta} = \frac{\partial \vartheta}{\partial R} \nabla R + \frac{\partial \vartheta}{\partial Z} \nabla Z \quad \mathbf{a}_{\vartheta} = \frac{\partial R}{\partial \vartheta} \mathbf{a}_R + \frac{\partial Z}{\partial \vartheta} \mathbf{a}_Z \quad (\text{D.27})$$

$$\nabla_{\zeta} = -\nabla \Phi \quad \mathbf{a}_{\zeta} = -\mathbf{a}_{\Phi} \quad (\text{D.28})$$

with metric components

$$g_{ij} = \begin{pmatrix} g_{\varrho\varrho} & g_{\varrho\vartheta} & 0 \\ g_{\varrho\vartheta} & g_{\vartheta\vartheta} & 0 \\ 0 & 0 & g_{\zeta\zeta} \end{pmatrix}, \quad g^{ij} = \begin{pmatrix} g^{\varrho\varrho} & g^{\varrho\vartheta} & 0 \\ g^{\varrho\vartheta} & g^{\vartheta\vartheta} & 0 \\ 0 & 0 & g^{\zeta\zeta} \end{pmatrix}. \quad (\text{D.29})$$

Explicitly, the covariant metric components are

$$g_{\varrho\varrho} = \left( \frac{\partial R}{\partial \varrho} \right)^2 + \left( \frac{\partial Z}{\partial \varrho} \right)^2 \quad (\text{D.30})$$

$$g_{\varrho\vartheta} = \frac{\partial R}{\partial \varrho} \frac{\partial R}{\partial \vartheta} + \frac{\partial Z}{\partial \varrho} \frac{\partial Z}{\partial \vartheta} \quad (\text{D.31})$$

$$g_{\vartheta\vartheta} = \left( \frac{\partial R}{\partial \vartheta} \right)^2 + \left( \frac{\partial Z}{\partial \vartheta} \right)^2 \quad (\text{D.32})$$

$$g_{\zeta\zeta} = R^2 \quad (\text{D.33})$$

where the contravariant metric terms can be obtained from equation C.50. Note that because  $g_{\varrho\zeta} = g_{\vartheta\zeta} = 0$ , we have

$$g^{\zeta\zeta} = 1/g_{\zeta\zeta} = 1/R^2. \quad (\text{D.34})$$

The Jacobian for the system is

$$\mathcal{J}_{\varrho\vartheta\zeta} = R \left( \frac{\partial R}{\partial \varrho} \frac{\partial Z}{\partial \vartheta} - \frac{\partial R}{\partial \vartheta} \frac{\partial Z}{\partial \varrho} \right). \quad (\text{D.35})$$

with radial derivative

$$\frac{\partial \mathcal{J}_{\varrho\vartheta\zeta}}{\partial \varrho} = \frac{\partial R}{\partial \varrho} \frac{\mathcal{J}_{\varrho\vartheta\zeta}}{R} + R \left( \frac{\partial^2 R}{\partial \varrho^2} \frac{\partial Z}{\partial \vartheta} + \frac{\partial R}{\partial \varrho} \frac{\partial^2 Z}{\partial \vartheta \partial \varrho} - \frac{\partial^2 R}{\partial \vartheta \partial \varrho} \frac{\partial Z}{\partial \varrho} - \frac{\partial R}{\partial \vartheta} \frac{\partial^2 Z}{\partial \varrho^2} \right). \quad (\text{D.36})$$

Note that all metric components are independent of  $\zeta$  as a consequence of axisymmetry.

### D.3.1 The potential $G$ in toroidal coordinates

The potential  $G$  in cylindrical coordinates is given by equation 3.111. Transforming this to toroidal coordinates gives  $G = \tilde{G}(\varrho, \vartheta) - \bar{c}_1(\varrho) \zeta$ . The form of  $G$  can be constrained by noting

that the equilibrium magnetic field is physical, and so its components must be periodic in  $\vartheta$ . Evaluating the gradient of  $G$  in  $\mathbf{B} = \nabla\varrho \times \nabla G$  (equation D.20), one finds

$$\begin{aligned}\mathbf{B} &= \frac{\partial G}{\partial \zeta} \nabla\varrho \times \nabla\zeta + \frac{\partial G}{\partial \vartheta} \nabla\varrho \times \nabla\vartheta \\ &= \frac{\bar{c}_1}{\mathcal{J}_{\varrho\vartheta\zeta}} \mathbf{a}_\vartheta + \frac{1}{\mathcal{J}_{\varrho\vartheta\zeta}} \frac{\partial \tilde{G}}{\partial \vartheta} \mathbf{a}_\zeta\end{aligned}\tag{D.37}$$

and so  $B^\zeta = (1/\mathcal{J}_{\varrho\vartheta\zeta}) \partial \tilde{G} / \partial \vartheta$ . Because  $B^\zeta$  and the Jacobian are periodic,  $\partial \tilde{G} / \partial \vartheta$  must itself be periodic. The integral of a periodic function produces a linear part and a periodic part<sup>1</sup>, and thus  $\tilde{G}$  has the form

$$\tilde{G}(\varrho, \vartheta) = \bar{c}_2(\varrho) \vartheta + \tilde{G}_0(\varrho, \vartheta)\tag{D.38}$$

where  $\tilde{G}_0(\varrho, \vartheta)$  is a function periodic in  $\vartheta$ . The potential  $G(\varrho, \vartheta, \zeta)$  is therefore

$$G = \bar{c}_2(\varrho) \vartheta - \bar{c}_1(\varrho) \zeta + \tilde{G}_0(\varrho, \vartheta).\tag{D.39}$$

### D.3.2 Magnetic fluxes

For the poloidal surface element  $d\mathbf{S}_\vartheta = \mathbf{a}_\zeta \times \mathbf{a}_\varrho d\varrho d\zeta = \mathcal{J}_{\varrho\vartheta\zeta} \nabla\vartheta d\varrho d\zeta$  and the toroidal surface element  $d\mathbf{S}_\zeta = \mathbf{a}_\varrho \times \mathbf{a}_\vartheta d\varrho d\vartheta = \mathcal{J}_{\varrho\vartheta\zeta} \nabla\zeta d\varrho d\vartheta$ , we define the poloidal flux and toroidal flux as

$$\Phi_\vartheta = \iint \mathbf{B} \cdot d\mathbf{S}_\vartheta, \quad \Phi_\zeta = \iint \mathbf{B} \cdot d\mathbf{S}_\zeta.\tag{D.40}$$

For the poloidal flux, using equation 3.113, one finds

$$\Phi_\vartheta(\varrho) = \int_0^\varrho \int_{-\pi}^\pi \bar{c}_1(\varrho') d\varrho' d\zeta\tag{D.41}$$

and therefore

$$\bar{c}_1(\varrho) = \frac{1}{2\pi} \frac{d\Phi_\vartheta}{d\varrho} = \psi'\tag{D.42}$$

where, due to the frequent occurrence of this quantity, the symbol  $\psi$  is introduced to denote the poloidal flux divided by  $2\pi$ , and the prime indicates its radial derivative. Similarly for the toroidal flux, one finds

$$\Phi_\zeta(\varrho) = \int_0^\varrho \int_{-\pi}^\pi \left( \bar{c}_2(\varrho') + \frac{\partial \tilde{G}_0}{\partial \vartheta} \right) d\varrho' d\vartheta = 2\pi \int_0^\varrho c_2(\varrho') d\varrho'\tag{D.43}$$

where we have used that  $\tilde{G}_0$  is periodic in  $\vartheta$ . Therefore we have

$$\bar{c}_2(\varrho) = \frac{1}{2\pi} \frac{d\Phi_\zeta}{d\varrho} = \chi'\tag{D.44}$$

defining the toroidal flux divided by  $2\pi$ ,  $\chi$ .

<sup>1</sup>Expressing a periodic function  $p(\varrho, \vartheta)$  as a Fourier series via  $p = \hat{p}_0(\varrho) + \sum_{n \neq 0} \hat{p}_n(\varrho) e^{in\vartheta}$ , its integral is  $\int p d\vartheta = \hat{p}_0(\varrho) \vartheta + \sum_{n \neq 0} (-i/n) \hat{p}_n(\varrho) e^{in\vartheta} + C(\varrho)$ , where  $C(\varrho) + \sum_{n \neq 0} (-i/n) \hat{p}_n(\varrho) e^{in\vartheta}$  is itself periodic.

### D.3.3 Alternative angular coordinates

For the toroidal coordinate system, the angles  $\vartheta$  and  $\zeta$  are not the only possible choices for angular coordinates. Any function that is periodic in  $\zeta$  at constant  $\vartheta$ , but increases by a period per poloidal turn at constant  $\zeta$ , is a viable alternative poloidal parameterisation of the flux surface. Such an alternative angle  $\vartheta_A$  with period  $\Delta\vartheta_A$  therefore satisfies

$$\vartheta_A(\varrho, \vartheta + 2M\pi, \zeta + 2N\pi) = \vartheta_A(\varrho, \vartheta, \zeta) + M\Delta\vartheta_A \quad (\text{D.45})$$

Similarly for an alternative toroidal angle,

$$\zeta_A(\varrho, \vartheta + 2M\pi, \zeta + 2N\pi) = \zeta_A(\varrho, \vartheta, \zeta) + N\Delta\zeta_A. \quad (\text{D.46})$$

The general solutions to equations D.45 and D.46 are of the form

$$\vartheta_A = \frac{\Delta\vartheta_A}{2\pi} (\vartheta + p_\vartheta(\varrho, \vartheta, \zeta)) \quad (\text{D.47})$$

$$\zeta_A = \frac{\Delta\zeta_A}{2\pi} (\zeta + p_\zeta(\varrho, \vartheta, \zeta)) \quad (\text{D.48})$$

where  $p_\vartheta$  and  $p_\zeta$  are functions periodic in both  $\vartheta$  and  $\zeta$ . The freedom in defining these periodic functions can be exploited to describe coordinates with useful properties.

We would like to keep axisymmetry associated with the toroidal coordinate only. Via the chain rule we have

$$\frac{\partial}{\partial \zeta} = \frac{\partial \vartheta_A}{\partial \zeta} \frac{\partial}{\partial \vartheta_A} + \frac{\partial \zeta_A}{\partial \zeta} \frac{\partial}{\partial \zeta_A} \quad (\text{D.49})$$

and thus we make the choice that  $\partial p_\vartheta / \partial \zeta = \partial p_\zeta / \partial \zeta = 0$ , giving<sup>2</sup>

$$\frac{\partial}{\partial \zeta} = \frac{\Delta\zeta_A}{2\pi} \frac{\partial}{\partial \zeta_A}. \quad (\text{D.50})$$

The transformations are therefore

$$\varrho_A = \varrho \qquad \varrho = \varrho_A \quad (\text{D.51})$$

$$\vartheta_A = \frac{\Delta\vartheta_A}{2\pi} (\vartheta + p_\vartheta(\varrho, \vartheta)) \qquad \vartheta = \frac{2\pi}{\Delta\vartheta_A} (\vartheta_A + p_{\vartheta_A}(\varrho_A, \vartheta_A)) \quad (\text{D.52})$$

$$\zeta_A = \frac{\Delta\zeta_A}{2\pi} (\zeta + p_\zeta(\varrho, \vartheta)) \qquad \zeta = \frac{2\pi}{\Delta\zeta_A} (\zeta_A + p_{\zeta_A}(\varrho_A, \vartheta_A)) \quad (\text{D.53})$$

---

<sup>2</sup>Strictly one does not require  $\partial p_\zeta / \partial \zeta = 0$  to maintain axisymmetry in  $\zeta_A$ , however given the equilibrium geometry is axisymmetric little generalisation is lost by this choice.

with basis vectors

$$\nabla_{\varrho_A} = \nabla_{\varrho} \qquad \mathbf{a}_{\varrho_A} = \mathbf{a}_{\varrho} + \frac{2\pi}{\Delta\vartheta_A} \frac{\partial p_{\vartheta_A}}{\partial \varrho_A} \mathbf{a}_{\vartheta} + \frac{2\pi}{\Delta\zeta_A} \frac{\partial p_{\zeta_A}}{\partial \varrho_A} \mathbf{a}_{\zeta} \quad (\text{D.54})$$

$$\nabla_{\vartheta_A} = \frac{\Delta\vartheta_A}{2\pi} \left[ \frac{\partial p_{\vartheta}}{\partial \varrho} \nabla_{\varrho} + \left( 1 + \frac{\partial p_{\vartheta}}{\partial \vartheta} \right) \nabla_{\vartheta} \right] \qquad \mathbf{a}_{\vartheta_A} = \frac{2\pi}{\Delta\vartheta_A} \left( 1 + \frac{\partial p_{\vartheta_A}}{\partial \vartheta_A} \right) \mathbf{a}_{\vartheta} + \frac{2\pi}{\Delta\zeta_A} \frac{\partial p_{\zeta_A}}{\partial \vartheta_A} \mathbf{a}_{\zeta} \quad (\text{D.55})$$

$$\nabla_{\zeta_A} = \frac{\Delta\zeta_A}{2\pi} \left[ \frac{\partial p_{\zeta}}{\partial \varrho} \nabla_{\varrho} + \frac{\partial p_{\zeta}}{\partial \vartheta} \nabla_{\vartheta} + \nabla_{\zeta} \right] \qquad \mathbf{a}_{\zeta_A} = \frac{2\pi}{\Delta\zeta_A} \mathbf{a}_{\zeta} \quad (\text{D.56})$$

from which the metric components can be obtained. Note that all components are generally non-zero, including  $g_{\varrho_A\zeta_A}$  and  $g_{\vartheta_A\zeta_A}$ .

### Freedom in $p_{\vartheta}$

Using the contravariant basis vectors to find the Jacobian of the alternative angle system,  $\mathcal{J}_A$ , one has

$$\begin{aligned} \frac{1}{\mathcal{J}_A} &= \nabla_{\varrho_A} \cdot \nabla_{\vartheta_A} \times \nabla_{\zeta_A} \\ &= \frac{\Delta\vartheta_A \Delta\zeta_A}{(2\pi)^2} \left( 1 + \frac{\partial p_{\vartheta}}{\partial \vartheta} \right) \frac{1}{\mathcal{J}_{\varrho\vartheta\zeta}} \end{aligned} \quad (\text{D.57})$$

and so the freedom in  $p_{\vartheta}$  can be used to set the angular dependence of the new Jacobian. Of particular interest is to make the new Jacobian a flux function,  $\mathcal{J}_A = \mathcal{J}_A(\varrho_A)$ . Solving equation D.57 in this case, specifying the periodicity of  $p_{\vartheta}$  through the condition  $p_{\vartheta}(\varrho, \vartheta_0 + 2\pi) = p_{\vartheta}(\varrho, \vartheta_0)$  gives

$$p_{\vartheta}(\varrho, \vartheta) = p_{\vartheta}(\varrho, \vartheta_0) + \frac{\int_{\vartheta_0}^{\vartheta} \mathcal{J}_{\varrho\vartheta\zeta} d\vartheta'}{\langle \mathcal{J}_{\varrho\vartheta\zeta} \rangle_{\text{P}}} - (\vartheta - \vartheta_0) \quad (\text{D.58})$$

where  $p_{\vartheta}(\varrho, \vartheta_0)$  is an integration constant, and

$$\mathcal{J}_A(\varrho_A) = \frac{(2\pi)^2}{\Delta\vartheta_A \Delta\zeta_A} \langle \mathcal{J}_{\varrho\vartheta\zeta} \rangle_{\text{P}} \quad (\text{D.59})$$

where  $\langle \dots \rangle_{\text{P}}$  denotes an average over  $\vartheta$ , defined by equation 3.133. Note that when the Jacobian is a flux function, the contravariant poloidal component of the magnetic field also becomes a flux function,  $B^{\vartheta_A} = 2\pi\psi' / (\mathcal{J}_A \Delta\zeta_A)$ .

Alternatively, one can use the freedom in  $p_{\vartheta}$  to construct a poloidal angle for which the coefficient of the spatial derivative parallel to the equilibrium magnetic field is a flux function. This is achieved by having a poloidal angle which satisfies

$$\hat{\mathbf{b}} \cdot \nabla_{\vartheta_A} = g_A(\varrho_A) \quad (\text{D.60})$$

for flux function  $g_A$ . In the toroidal system we have  $\hat{\mathbf{b}} = |\mathbf{B}|^{-1} (B^\vartheta \mathbf{a}_\vartheta + B^\zeta \mathbf{a}_\zeta)$  which gives

$$\frac{B^\vartheta}{|\mathbf{B}|} \frac{\Delta\vartheta_A}{2\pi} \left( 1 + \frac{\partial p_\vartheta}{\partial \vartheta} \right) = g_A. \quad (\text{D.61})$$

This integrates to, again specifying the periodicity of  $p_\vartheta$ ,

$$p_\vartheta(\varrho, \vartheta) = p_\vartheta(\varrho, \vartheta_0) + \frac{\int_{\vartheta_0}^{\vartheta} \frac{|\mathbf{B}|}{B^\vartheta} d\vartheta'}{\left\langle \frac{|\mathbf{B}|}{B^\vartheta} \right\rangle_{\text{P}}} - (\vartheta - \vartheta_0) \quad (\text{D.62})$$

where  $|\mathbf{B}|/B^\vartheta = (1/B^\vartheta) \sqrt{B_\vartheta B^\vartheta + B_\zeta B^\zeta} = \sqrt{g_{\vartheta\vartheta} + q^2 (\mathcal{J}_{\vartheta\vartheta\zeta})^2 g^{\zeta\zeta} / \langle \mathcal{J}_{\vartheta\vartheta\zeta} g^{\zeta\zeta} \rangle_{\text{P}}^2}$ , and

$$g_A = \frac{\Delta\vartheta_A}{2\pi} \frac{1}{\left\langle \frac{|\mathbf{B}|}{B^\vartheta} \right\rangle_{\text{P}}}. \quad (\text{D.63})$$

The Jacobian under this choice is

$$\mathcal{J}_A = \frac{(2\pi)^2}{\Delta\vartheta_A \Delta\zeta_A} \mathcal{J}_{\varrho\vartheta\zeta} \frac{\left\langle \frac{|\mathbf{B}|}{B^\vartheta} \right\rangle_{\text{P}}}{|\mathbf{B}|/B^\vartheta}. \quad (\text{D.64})$$

### Freedom in $p_\zeta$

Consider the form of the second potential that describes the equilibrium magnetic field,  $q(\varrho)\vartheta - \zeta + G_0(\varrho, \vartheta)$  (equation 3.115) where  $G_0$  is given by equation 3.131. Substituting the alternative angles, the potential becomes

$$q\vartheta - \zeta + G_0 = \frac{2\pi}{\Delta\vartheta_A} q\vartheta_A - \frac{2\pi}{\Delta\zeta_A} \zeta_A + p_\zeta - qp_\vartheta + G_0. \quad (\text{D.65})$$

If we make the choice that  $p_\zeta = qp_\vartheta - G_0$ , the potential loses its periodic part due to it being absorbed into the definition of  $\zeta_A$ , regardless of our choice of  $p_\vartheta$ . In this case, equation 3.115 is

$$\begin{aligned} \mathbf{B} &= \psi' \nabla_{\varrho_A} \times \nabla \left[ \frac{2\pi}{\Delta\vartheta_A} q(\varrho_A) \vartheta_A - \frac{2\pi}{\Delta\zeta_A} \zeta_A \right] \\ &= \frac{2\pi}{\Delta\zeta_A} \frac{\psi'}{\mathcal{J}_A} \mathbf{a}_{\vartheta_A} + q \frac{2\pi}{\Delta\vartheta_A} \frac{\psi'}{\mathcal{J}_A} \mathbf{a}_{\zeta_A} \end{aligned} \quad (\text{D.66})$$

and so the ratio of the contravariant components becomes a flux function,

$$\frac{B^{\zeta_A}}{B^{\vartheta_A}} = q \frac{\Delta\zeta_A}{\Delta\vartheta_A} \quad (\text{D.67})$$

Note therefore that if one makes this choice together with the form of  $p_\vartheta$  that makes the Jacobian a flux function, then both contravariant components of the equilibrium magnetic field become flux functions. These are Hamada coordinates.

We again emphasise care should be taken when comparing results between gyrokinetic codes, as the definitions of the angular coordinates can differ. For example, evenly spaced  $\vartheta$  grids do not necessarily correspond to equally spaced  $\vartheta_A$  grids.

Later when transforming to field-aligned coordinates, as done in section 3.4.8, one can define analogous field-aligned coordinates based on these alternative angles, i.e  $r_A = \varrho_A$ ,  $\alpha_A = \frac{2\pi}{\Delta\vartheta_A}q\vartheta_A - \frac{2\pi}{\Delta\zeta_A}\zeta_A + p_\zeta - qp\vartheta + G_0$ ,  $\theta_A = \vartheta_A$ . The properties of the alternative angles (for example a flux function Jacobian) are then carried forward.

### GKW coordinates

As an example we show the coordinates used by GKW [35], which are Hamada coordinates defined with a unit-period,  $\Delta\vartheta_{\text{GKW}} = \Delta\zeta_{\text{GKW}} = 1$ . The transformations are

$$\varrho_{\text{GKW}} = \varrho \quad (\text{D.68})$$

$$\vartheta_{\text{GKW}} = \frac{1}{2\pi} \left( \vartheta + \left[ \frac{\int_0^\vartheta \mathcal{J}_{\varrho\vartheta\zeta} d\vartheta'}{\langle \mathcal{J}_{\varrho\vartheta\zeta} \rangle_{\text{P}}} - \vartheta \right] \right) \quad (\text{D.69})$$

$$\zeta_{\text{GKW}} = \frac{1}{2\pi} \left( \zeta + q \left[ \frac{\int_0^\vartheta \mathcal{J}_{\varrho\vartheta\zeta} d\vartheta'}{\langle \mathcal{J}_{\varrho\vartheta\zeta} \rangle_{\text{P}}} - \vartheta \right] - q \left[ \frac{\int_0^\vartheta \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} d\vartheta'}{\langle \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} \rangle_{\text{P}}} - \vartheta \right] \right) \quad (\text{D.70})$$

where the integration constants have been set such that  $\vartheta_{\text{GKW}}(\varrho, 0) = 0$ ,  $\zeta_{\text{GKW}}(\varrho, 0, 0) = 0$ , and the quantities enclosed by square brackets are periodic. The Jacobian is

$$\mathcal{J}_{\text{GKW}} = (2\pi)^2 \langle \mathcal{J}_{\varrho\vartheta\zeta} \rangle_{\text{P}}. \quad (\text{D.71})$$

### D.3.4 The standard Grad-Shafranov equation

Starting from equation 3.134, one uses equation 3.128 for  $\mu_0 J^\vartheta$ ,  $B^\vartheta = \psi' / \mathcal{J}_{\varrho\vartheta\zeta}$ ,  $B^\zeta = g^{\zeta\zeta} B_\zeta$  and the substitution  $J^\zeta = -J^\Phi$  to find

$$\mu_0 J^\Phi \psi' = \mu_0 \frac{dP}{d\varrho} + g^{\zeta\zeta} B_\zeta \frac{dB_\zeta}{d\varrho}. \quad (\text{D.72})$$

As  $\psi'$  is a flux function, dividing both sides by  $\psi'$  allows the total derivatives to be written with respect to  $\psi$ ,

$$\mu_0 J^\Phi = \mu_0 \frac{dP}{d\psi} + g^{\zeta\zeta} B_\zeta \frac{dB_\zeta}{d\psi}. \quad (\text{D.73})$$

The form of  $\mu_0 J^\Phi$  can be found using Ampère's law, equation 3.126, now evaluated in the cylindrical system

$$\begin{aligned} \mu_0 J^\Phi &= \frac{1}{R} \left( \frac{\partial B_R}{\partial Z} - \frac{\partial B_Z}{\partial R} \right) \\ &= \frac{1}{R} \left( \frac{\partial B^R}{\partial Z} - \frac{\partial B^Z}{\partial R} \right) \end{aligned} \quad (\text{D.74})$$

where we have used  $B_R = B^R$  and  $B_Z = B^Z$ . Using equations D.16 and D.18, where it has previously been shown  $df_p/d\varrho = \bar{c}_1 = \psi'$  (equation D.42),  $\mu_0 J^\Phi$  becomes

$$\mu_0 J^\Phi = -\frac{1}{R^2} \left[ \frac{\partial^2 \psi}{\partial R^2} - \frac{1}{R} \frac{\partial \psi}{\partial R} + \frac{\partial^2 \psi}{\partial Z^2} \right]. \quad (\text{D.75})$$

Substituting this into equation D.73, multiplying both sides by  $-R^2$  and using  $g^{\zeta\zeta} = 1/R^2$  gives the Grad-Shafranov equation,

$$\frac{\partial^2 \psi}{\partial R^2} - \frac{1}{R} \frac{\partial \psi}{\partial R} + \frac{\partial^2 \psi}{\partial Z^2} = -\mu_0 \frac{dP}{d\psi} R^2 - B_\zeta \frac{dB_\zeta}{d\psi}. \quad (\text{D.76})$$

## D.4 Global field-aligned coordinates

### D.4.1 Third coordinate considerations

Let us initially label the third coordinate of the field-aligned system with the symbol  $\varsigma$ . Given that the covariant basis vector of  $\varsigma$  is aligned with the magnetic field regardless of the transformation  $\varsigma = \varsigma(\varrho, \vartheta, \zeta)$  from equation 3.109, a choice must be made as to its form. We consider several desirable properties of the new system to constrain it.

Considering axisymmetry, we wish to continue to restrict it to a single coordinate. Using the chain rule, we find

$$\begin{aligned} \frac{\partial}{\partial \zeta} &= \frac{\partial r}{\partial \zeta} \frac{\partial}{\partial r} + \frac{\partial \alpha}{\partial \zeta} \frac{\partial}{\partial \alpha} + \frac{\partial \varsigma}{\partial \zeta} \frac{\partial}{\partial \varsigma} \\ &= -\frac{\partial}{\partial \alpha} + \frac{\partial \varsigma}{\partial \zeta} \frac{\partial}{\partial \varsigma} \end{aligned} \quad (\text{D.77})$$

and so to isolate the effect of axisymmetry to the  $\alpha$  coordinate, we choose  $\partial \varsigma / \partial \zeta = 0$ .

Using equation C.26, the inverse of the new Jacobian is

$$\begin{aligned} \frac{1}{\mathcal{J}_{r\alpha\varsigma}} &= \nabla r \cdot (\nabla \alpha \times \nabla \varsigma) \\ &= \nabla \varrho \cdot (\nabla \zeta \times \nabla \vartheta) \frac{\partial \alpha}{\partial \zeta} \frac{\partial \varsigma}{\partial \vartheta} \\ &= \frac{1}{\mathcal{J}_{\varrho\vartheta\zeta}} \frac{\partial \varsigma}{\partial \vartheta} \end{aligned} \quad (\text{D.78})$$

and so the Jacobian remains unchanged provided  $\partial \varsigma / \partial \vartheta = 1$ . By choosing  $\varsigma = \vartheta$ , the formulation for the flux-surface parameterisation can be trivially carried over to the new system, for example  $R(\varrho_0, \vartheta) \rightarrow R(r_0, \varsigma)$ .

For these three reasons we choose  $\varsigma = \vartheta$ , and thus relabel the third field-aligned coordinate  $\varsigma \rightarrow \theta = \vartheta$ .

## D.4.2 Geometric quantities

### Basis vectors and metric components

Using the transformations in section 3.4.8, the basis vectors in the global field-aligned system are

$$\nabla r = \nabla \varrho \qquad \mathbf{a}_r = \mathbf{a}_\varrho + \left[ \frac{dq}{dr} \theta + \frac{\partial G_0}{\partial r} \right] \mathbf{a}_\zeta \quad (\text{D.79})$$

$$\nabla \alpha = \left[ \frac{dq}{d\varrho} \vartheta + \frac{\partial G_0}{\partial \varrho} \right] \nabla \varrho + \left[ q + \frac{\partial G_0}{\partial \vartheta} \right] \nabla \vartheta - \nabla \zeta \qquad \mathbf{a}_\alpha = -\mathbf{a}_\zeta \quad (\text{D.80})$$

$$\nabla \theta = \nabla \vartheta \qquad \mathbf{a}_\theta = \mathbf{a}_\vartheta + \left[ q(r) + \frac{\partial G_0}{\partial \theta} \right] \mathbf{a}_\zeta \quad (\text{D.81})$$

with covariant metric components

$$g_{rr} = g_{\varrho\varrho} + \left[ \frac{dq}{dr} \theta + \frac{\partial G_0}{\partial r} \right]^2 g_{\zeta\zeta} \quad (\text{D.82})$$

$$g_{r\alpha} = - \left[ \frac{dq}{dr} \theta + \frac{\partial G_0}{\partial r} \right] g_{\zeta\zeta} \quad (\text{D.83})$$

$$g_{r\theta} = g_{\varrho\vartheta} + \left[ \frac{dq}{dr} \theta + \frac{\partial G_0}{\partial r} \right] \left[ q(r) + \frac{\partial G_0}{\partial \theta} \right] g_{\zeta\zeta} \quad (\text{D.84})$$

$$g_{\alpha\alpha} = g_{\zeta\zeta} \quad (\text{D.85})$$

$$g_{\alpha\theta} = - \left[ q(r) + \frac{\partial G_0}{\partial \theta} \right] g_{\zeta\zeta} \quad (\text{D.86})$$

$$g_{\theta\theta} = g_{\vartheta\vartheta} + \left[ q(r) + \frac{\partial G_0}{\partial \theta} \right]^2 g_{\zeta\zeta} \quad (\text{D.87})$$

from which one can calculate the contravariant components using C.50.

The partial derivatives of the toroidal system in terms of the global field-aligned are

$$\left. \frac{\partial}{\partial \zeta} \right|_{\varrho, \vartheta} = - \left. \frac{\partial}{\partial \alpha} \right|_{r, \theta} \quad (\text{D.88})$$

$$\left. \frac{\partial}{\partial \vartheta} \right|_{\zeta, \varrho} = \left. \frac{\partial \alpha}{\partial \vartheta} \right|_{\zeta, \varrho} \left. \frac{\partial}{\partial \alpha} \right|_{r, \theta} + \left. \frac{\partial}{\partial \theta} \right|_{r, \alpha} \quad (\text{D.89})$$

$$\left. \frac{\partial}{\partial \varrho} \right|_{\vartheta, \zeta} = \left. \frac{\partial \alpha}{\partial \varrho} \right|_{\vartheta, \zeta} \left. \frac{\partial}{\partial \alpha} \right|_{r, \theta} + \left. \frac{\partial}{\partial r} \right|_{\alpha, \theta} \quad (\text{D.90})$$

and thus we see that for functions that are axisymmetric, one has

$$\frac{\partial}{\partial \varrho} = \frac{\partial}{\partial r}, \quad \frac{\partial}{\partial \vartheta} = \frac{\partial}{\partial \theta} \quad (\text{Axisymmetric quantities}) \quad (\text{D.91})$$

The derivatives of the two coordinate systems may be used interchangeably in this case, such as with  $G_0$ .

To evaluate the square brackets in the definitions of the basis vectors and metric components, we use equation 3.131 to obtain

$$q + \frac{\partial G_0}{\partial \vartheta} = q \frac{\mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta}}{\langle \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} \rangle_{\text{P}}} = \frac{B_\zeta}{\psi'} \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta}. \quad (\text{D.92})$$

The square bracket containing radial derivatives can be calculated by differentiating equation D.92 with respect to  $\varrho$ , then integrating over  $\vartheta$ . Differentiating, we get

$$\frac{dq}{d\varrho} + \frac{\partial^2 G_0}{\partial \varrho \partial \vartheta} = \frac{dB_\zeta}{d\varrho} \frac{\mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta}}{\psi'} - \frac{\psi''}{(\psi')^2} \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} B_\zeta + \frac{B_\zeta}{\psi'} \frac{\partial \mathcal{J}_{\varrho\vartheta\zeta}}{\partial \varrho} g^{\zeta\zeta} + \frac{B_\zeta}{\psi'} \frac{\partial g^{\zeta\zeta}}{\partial \varrho} \mathcal{J}_{\varrho\vartheta\zeta}. \quad (\text{D.93})$$

and now integrating over  $\vartheta$ , and using equation 3.137, one finds

$$\begin{aligned} \frac{dq}{d\varrho} \vartheta + \frac{\partial G_0}{\partial \varrho} &= \frac{dq}{d\varrho} \vartheta_0 + \left. \frac{\partial G_0}{\partial \varrho} \right|_{\vartheta_0} \\ &+ \frac{1}{\psi'} \frac{dB_\zeta}{d\varrho} \int_{\vartheta_0}^{\vartheta} \left( \mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta} + \left[ \frac{B_\zeta}{\psi'} \right]^2 \frac{(\mathcal{J}_{\varrho\vartheta\zeta})^3 (g^{\zeta\zeta})^2}{g_{\vartheta\vartheta}} \right) d\vartheta' \\ &+ \frac{B_\zeta}{\psi'} \int_{\vartheta_0}^{\vartheta} \left( \frac{(\mathcal{J}_{\varrho\vartheta\zeta})^3 g^{\zeta\zeta}}{g_{\vartheta\vartheta}} \left[ \frac{\mu_0}{(\psi')^2} \frac{dP}{d\varrho} \right] + \mathcal{J}_{\varrho\vartheta\zeta} \frac{\partial g^{\zeta\zeta}}{\partial \varrho} \right) d\vartheta' \\ &- \frac{B_\zeta}{\psi'} \int_{\vartheta_0}^{\vartheta} \frac{\mathcal{J}_{\varrho\vartheta\zeta} g^{\zeta\zeta}}{g_{\vartheta\vartheta}} \left( \frac{\partial g_{\vartheta\vartheta}}{\partial \vartheta} - \frac{\partial g_{\vartheta\vartheta}}{\partial \varrho} - \frac{g_{\vartheta\vartheta}}{\mathcal{J}_{\varrho\vartheta\zeta}} \frac{\partial \mathcal{J}_{\varrho\vartheta\zeta}}{\partial \vartheta} \right) d\vartheta' \end{aligned} \quad (\text{D.94})$$

where  $dB_\zeta/d\varrho$  is given by equation 3.139.

### Equilibrium magnetic field gradient

In the global field-aligned system, one has  $\nabla B = \partial B/\partial r \nabla r + \partial B/\partial \theta \nabla \theta$ . The partial derivatives are, using D.91,

$$\begin{aligned} \frac{\partial B}{\partial r} &= \frac{1}{2B} \frac{\partial (B^2)}{\partial r} \\ &= \frac{1}{2B} \frac{\partial (B_\vartheta B^\vartheta + B_\zeta B^\zeta)}{\partial r} \\ &= \frac{1}{2B} \frac{\partial (g_{\vartheta\vartheta} (B^\vartheta)^2 + g^{\zeta\zeta} (B_\zeta)^2)}{\partial r} \\ &= \frac{1}{2\sqrt{g_{\vartheta\vartheta} (B^\vartheta)^2 + g^{\zeta\zeta} (B_\zeta)^2}} \left[ \frac{\partial g_{\vartheta\vartheta}}{\partial \varrho} (B^\vartheta)^2 + 2g_{\vartheta\vartheta} B^\vartheta \frac{\partial B^\vartheta}{\partial \varrho} + \frac{\partial g^{\zeta\zeta}}{\partial \varrho} (B_\zeta)^2 + 2g^{\zeta\zeta} B_\zeta \frac{dB_\zeta}{d\varrho} \right] \end{aligned} \quad (\text{D.95})$$

where  $\partial B^\vartheta/\partial \varrho = \psi''/\mathcal{J}_{\varrho\vartheta\zeta} - (\psi'/\mathcal{J}_{\varrho\vartheta\zeta})^2 (\partial \mathcal{J}_{\varrho\vartheta\zeta}/\partial \varrho)$ , with  $\partial \mathcal{J}_{\varrho\vartheta\zeta}/\partial \varrho$  given by equation 3.137, and  $dB_\zeta/d\varrho$  given by equation 3.139. Similarly for the  $\theta$  derivative, one finds

$$\frac{\partial B}{\partial \theta} = \frac{1}{2\sqrt{g_{\vartheta\vartheta} (B^\vartheta)^2 + g^{\zeta\zeta} (B_\zeta)^2}} \left[ \frac{\partial g_{\vartheta\vartheta}}{\partial \vartheta} (B^\vartheta)^2 + 2g_{\vartheta\vartheta} B^\vartheta \frac{\partial B^\vartheta}{\partial \vartheta} + \frac{\partial g^{\zeta\zeta}}{\partial \vartheta} (B_\zeta)^2 \right]. \quad (\text{D.96})$$

### The contravariant metric term $g^{\alpha\alpha}$

Here we explicitly show the form of  $g^{\alpha\alpha}$ , which is used to compare the agreement of the formalisms discussed in section 3.4.5 due to its dependence on the radial derivatives of the Jacobian and  $B_\zeta$ . Using equation C.50, it is given by

$$\begin{aligned} g^{\alpha\alpha} &= \frac{1}{(\mathcal{J}_{r\alpha\theta})^2} \left( g_{rr}g_{\theta\theta} - (g_{r\theta})^2 \right) \\ &= g^{\zeta\zeta} + \frac{\left[ q + \frac{\partial G_0}{\partial \vartheta} \right]^2 g_{\varrho\varrho} - 2 \left[ q + \frac{\partial G_0}{\partial \vartheta} \right] \left[ \frac{dq}{d\varrho} \vartheta + \frac{\partial G_0}{\partial \varrho} \right] g_{\varrho\vartheta} + \left[ \frac{dq}{d\varrho} \vartheta + \frac{\partial G_0}{\partial \varrho} \right]^2 g_{\vartheta\vartheta}}{g_{\varrho\varrho}g_{\vartheta\vartheta} - (g_{\varrho\vartheta})^2} \end{aligned} \quad (\text{D.97})$$

where the terms in the square brackets are given by equations D.92 and D.94.

## D.5 Local field-aligned system

### D.5.1 Geometric quantities

Transforming to the local field-aligned system from the global via  $\{r, \alpha, \theta\} \rightarrow \{x, y, \theta'\}$ , where  $x = r - r_0$ ,  $y = C_y \alpha$  and  $\theta = \theta'$ , we get

$$\nabla x = \nabla r \qquad \mathbf{a}_x = \mathbf{a}_r \quad (\text{D.98})$$

$$\nabla y = C_y \nabla \alpha \qquad \mathbf{a}_y = \frac{1}{C_y} \mathbf{a}_\alpha \quad (\text{D.99})$$

$$\nabla \theta' = \nabla \theta \qquad \mathbf{a}_{\theta'} = \mathbf{a}_\theta \quad (\text{D.100})$$

with covariant metric components  $g_{xx} = g_{rr}$ ,  $g_{xy} = g_{r\alpha}/C_y$ ,  $g_{x\theta'} = g_{r\theta}$ ,  $g_{yy} = g_{\alpha\alpha}/C_y^2$ ,  $g_{y\theta'} = g_{\alpha\theta}/C_y$  and  $g_{\theta'\theta'} = g_{\theta\theta}$ . Due to the equivalence of the basis vectors of  $\theta$  and  $\theta'$ , the prime can be dropped in the local field-aligned system. The Jacobian is  $\mathcal{J}_{xy\theta} = \mathcal{J}_{r\alpha\theta}/C_y$ , and using equation 3.189, for  $C_y = r_0/q(r_0)$  the magnetic field strength is  $B = B_{\text{unit}} \sqrt{g_{\theta\theta}}/\mathcal{J}_{xy\theta}$ . The radial partial derivative is related to the global system via

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial r} \quad (\text{D.101})$$

and thus can be equated with the radial derivatives in equation D.91 in the case of an axisymmetric quantity.

### D.5.2 Terms in the local gyrokinetic equation

Here we examine how the terms equation 3.97 appear in the local limit. Due to the frequency of the occurrence of vectors of the form  $\hat{\mathbf{b}} \times \nabla f$  for scalar function  $f$ , we write for reference

$$\begin{aligned}\hat{\mathbf{b}} \times \nabla f &= \left( \frac{1}{\sqrt{g_{\theta\theta}}} \mathbf{a}_\theta \right) \times \left( \frac{\partial f}{\partial \xi^k} \nabla \xi^k \right) \\ &= \frac{g_{\theta j}}{\sqrt{g_{\theta\theta}}} \frac{\partial f}{\partial \xi^k} \nabla \xi^j \times \nabla \xi^k \\ &= \frac{1}{\mathcal{J}_{xy\theta}} \epsilon^{ijk} \frac{g_{\theta j}}{\sqrt{g_{\theta\theta}}} \frac{\partial f}{\partial \xi^k} \mathbf{a}_i\end{aligned}\tag{D.102}$$

#### Parallel velocity

The parallel velocity term is

$$\begin{aligned}U_{0,\parallel} \hat{\mathbf{b}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} &= U_{0,\parallel} \left( \frac{1}{\sqrt{g_{\theta\theta}}} \mathbf{a}_\theta \right) \cdot \frac{\partial h_s}{\partial \xi^j} \nabla \xi^j \\ &= \frac{U_{0,\parallel}}{\sqrt{g_{\theta\theta}}} \frac{\partial h_s}{\partial \theta}\end{aligned}\tag{D.103}$$

#### Equilibrium magnetic drifts

The curvature drift may be rewritten using Ampère's law and force balance. Taking the cross product of Ampère's law with  $\mathbf{B}$ , and substituting force balance, gives

$$\begin{aligned}\mu_0 \nabla P &= -\mathbf{B} \times (\nabla \times \mathbf{B}) \\ &= (\mathbf{B} \cdot \nabla) \mathbf{B} - \frac{1}{2} \nabla B^2\end{aligned}\tag{D.104}$$

where we have used a vector identity. Expressing the magnetic field via  $\mathbf{B} = B \hat{\mathbf{b}}$  and using the chain rule, we get

$$\mu_0 \nabla P = B^2 \left( \hat{\mathbf{b}} \cdot \nabla \right) \hat{\mathbf{b}} + \hat{\mathbf{b}} (\mathbf{B} \cdot \nabla) B - \frac{1}{2} \nabla B^2.\tag{D.105}$$

Taking the cross product with  $\hat{\mathbf{b}}$  and rearranging, one finds

$$\hat{\mathbf{b}} \times \left[ \left( \hat{\mathbf{b}} \cdot \nabla \right) \hat{\mathbf{b}} \right] = \frac{\hat{\mathbf{b}} \times \nabla B}{B} + \frac{\hat{\mathbf{b}} \times \nabla P}{B^2/\mu_0}.\tag{D.106}$$

The equilibrium drift velocity becomes

$$\mathbf{v}_D = \frac{\left( U_{0,\parallel}^2 + u_0^2/2 \right)}{\Omega_s B} \hat{\mathbf{b}} \times \nabla B + \frac{U_{0,\parallel}^2}{\Omega_s} \frac{\hat{\mathbf{b}} \times \nabla P}{B^2/\mu_0}.\tag{D.107}$$

Note that this is also true globally. Now applied to the local limit, the product of the drift velocity with the gradient of  $h_s$  to lowest order is therefore, using equation D.102,

$$\begin{aligned} \mathbf{v}_D \cdot \frac{\partial h_s}{\partial \mathbf{R}} &= \frac{\left( U_{0,\parallel}^2 + u_0^2/2 \right)}{\Omega_s B_{\text{unit}}} \frac{g_{y\theta}}{g_{\theta\theta}} \frac{\partial B}{\partial \theta} \frac{\partial h_s}{\partial x} + \\ &\frac{\left( U_{0,\parallel}^2 + u_0^2/2 \right)}{\Omega_s B_{\text{unit}}} \left( \frac{\partial B}{\partial r} - \frac{g_{x\theta}}{g_{\theta\theta}} \frac{\partial B}{\partial \theta} \right) \frac{\partial h_s}{\partial y} + \\ &\frac{U_{0,\parallel}^2}{\Omega_s B_{\text{unit}}^2} \frac{\mathcal{J}_{xy\theta} \mu_0}{\sqrt{g_{\theta\theta}}} \frac{dP}{dr} \frac{\partial h_s}{\partial y}. \end{aligned} \quad (\text{D.108})$$

where  $\partial B/\partial r$  and  $\partial B/\partial \theta$  are given by equations D.95 and D.96 respectively.

### Fluctuating $E \times B$ velocity terms

Using equation D.102, the fluctuating  $E \times B$  velocity in the local field-aligned system is

$$\begin{aligned} \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} &= \frac{1}{\mathcal{J}_{xy\theta}} \epsilon^{ijk} \frac{g_{\theta j}}{B \sqrt{g_{\theta\theta}}} \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial \xi^k} \mathbf{a}_i \\ &= \epsilon^{ijk} \frac{1}{B_{\text{unit}}} \frac{g_{\theta j}}{g_{\theta\theta}} \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial \xi^k} \mathbf{a}_i \end{aligned} \quad (\text{D.109})$$

where we have used  $B = B_{\text{unit}} \sqrt{g_{\theta\theta}} / \mathcal{J}_{xy\theta}$ . Taking the dot product with the derivative of the Maxwellian, we get

$$\langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial F_{s,0}}{\partial \mathbf{R}} = - \frac{1}{B_{\text{unit}}} \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial y} F_{0s} \left[ \frac{1}{n_{0s}} \frac{dn_{0s}}{dr} + \frac{1}{T_{0s}} \frac{dT_{0s}}{dr} \left( \frac{\epsilon}{T_{0s}} - \frac{3}{2} \right) \right]. \quad (\text{D.110})$$

Now calculating the  $E \times B$  velocity acting on the gradient of  $h_s$ , one finds

$$\delta \left[ \langle \delta \mathbf{v}_E \rangle_{\mathbf{R}} \cdot \frac{\partial h_s}{\partial \mathbf{R}} \right] = \frac{1}{B_{\text{unit}}} \delta \left[ \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial x} \frac{\partial h_s}{\partial y} - \frac{\partial \langle \delta \phi \rangle_{\mathbf{R}}}{\partial y} \frac{\partial h_s}{\partial x} \right]. \quad (\text{D.111})$$

### D.5.3 Gyroaverage operators

We wish to explicitly calculate the gyroaverages in the local field-aligned system. Starting with the gyroaverage at constant  $\mathbf{R}$ , we consider the fluctuating potential,  $\delta\phi$ . The potential is evaluated at a position in space of  $\mathbf{R} + \boldsymbol{\rho}_0(t, \mathbf{R}, v_1, v_2, \varphi)$  from equation 3.26. It is the coordinates that must be considered at these positions, and so here we explicitly define where they are evaluated, such that  $x(\mathbf{R} + \boldsymbol{\rho}_0)$ ,  $y(\mathbf{R} + \boldsymbol{\rho}_0)$ , which differs from those coordinates at the gyrocentre position,  $x(\mathbf{R})$ ,  $y(\mathbf{R})$ . Using equation 3.191, the fluctuating potential in the local limit may be written

$$\delta\phi(\mathbf{R} + \boldsymbol{\rho}_0, t) = \sum_{k_x, k_y} \delta\hat{\phi}_{k_x, k_y}(\theta(\mathbf{R} + \boldsymbol{\rho}_0), t) e^{ik_x x(\mathbf{R} + \boldsymbol{\rho}_0)} e^{ik_y y(\mathbf{R} + \boldsymbol{\rho}_0)}. \quad (\text{D.112})$$

We now expand the coordinates around  $\mathbf{R}$ , and therefore the potential, to lowest order. Expanding the parallel coordinate,  $\theta(\mathbf{R} + \boldsymbol{\rho}_0) = \theta(\mathbf{R}) + (\boldsymbol{\rho}_0 \cdot \nabla)\theta|_{\mathbf{R}} + \mathcal{O}(\rho_*^2\theta)$ , the potential mode becomes

$$\begin{aligned} \delta\hat{\phi}_{k_x, k_y}(\theta(\mathbf{R} + \boldsymbol{\rho}_0), t) &= \delta\hat{\phi}_{k_x, k_y}(\theta(\mathbf{R}), t) + \boldsymbol{\rho}_0 \cdot \nabla \theta \frac{\partial \delta\hat{\phi}_{k_x, k_y}}{\partial \theta} + \dots \\ &= \delta\hat{\phi}_{k_x, k_y}(\theta(\mathbf{R}), t) + \mathcal{O}(\rho_* \delta\phi) \end{aligned} \quad (\text{D.113})$$

and so to lowest order the potential is simply evaluated at  $\theta(\mathbf{R})$ . Now we consider the exponentials of equation D.112. The combined argument is

$$\begin{aligned} k_x x(\mathbf{R} + \boldsymbol{\rho}_0) + k_y y(\mathbf{R} + \boldsymbol{\rho}_0) &= k_x x(\mathbf{R}) + k_y y(\mathbf{R}) + k_x \boldsymbol{\rho}_0 \cdot \nabla x + k_y \boldsymbol{\rho}_0 \cdot \nabla y + \dots \\ &= k_x x(\mathbf{R}) + k_y y(\mathbf{R}) + \boldsymbol{\rho}_0 \cdot (k_x \nabla x + k_y \nabla y) + \mathcal{O}(\rho_* \rho k_{\perp}) \end{aligned} \quad (\text{D.114})$$

and so we see the bracketed term is retained, as  $k_{\perp} \rho_0 \sim \mathcal{O}(1)$ . Equation D.112 becomes

$$\delta\phi(\mathbf{R} + \boldsymbol{\rho}_0, t) = \sum_{k_x, k_y} \delta\hat{\phi}_{k_x, k_y}(\theta(\mathbf{R}), t) e^{ik_x x(\mathbf{R})} e^{ik_y y(\mathbf{R})} e^{i\boldsymbol{\rho}_0 \cdot \mathbf{k}_{\perp}} + \mathcal{O}(\rho_* \delta\phi) \quad (\text{D.115})$$

where  $\mathbf{k}_{\perp} = k_x \nabla x + k_y \nabla y$  to lowest order. To evaluate the gyroaverage, we write the wavevector and gyroradius using  $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2$ . Defining an angle  $\bar{\chi}$ , such that  $\mathbf{k}_{\perp} = |\mathbf{k}_{\perp}| (\sin \bar{\chi} \hat{\mathbf{e}}_1 + \cos \bar{\chi} \hat{\mathbf{e}}_2)$ , and using  $\boldsymbol{\rho}_0 = \frac{u_0}{\Omega_s} (\sin \varphi \hat{\mathbf{e}}_1 + \cos \varphi \hat{\mathbf{e}}_2)$ , their scalar product becomes

$$\begin{aligned} \boldsymbol{\rho}_0 \cdot \mathbf{k}_{\perp} &= \frac{u_0 |\mathbf{k}_{\perp}|}{\Omega_s} (\sin \bar{\chi} \sin \varphi + \cos \bar{\chi} \cos \varphi) \\ &= \frac{u_0 |\mathbf{k}_{\perp}|}{\Omega_s} \cos(\bar{\chi} - \varphi) \end{aligned} \quad (\text{D.116})$$

where we have used a trigonometric identity. Upon gyroaveraging equation D.115, we get

$$\begin{aligned} \langle e^{i\boldsymbol{\rho}_0 \cdot \mathbf{k}_{\perp}} \rangle_{\mathbf{R}} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i \frac{u_0 |\mathbf{k}_{\perp}|}{\Omega} \cos(\chi - \varphi)} d\varphi \\ &= J_0 \left( \frac{u_0 |\mathbf{k}_{\perp}|}{\Omega} \right) \end{aligned} \quad (\text{D.117})$$

where  $J_0$  is the Bessel function of the first kind [159]. Therefore we have

$$\langle \delta\phi(\mathbf{R} + \boldsymbol{\rho}_0, t) \rangle_{\mathbf{R}} = \sum_{k_x, k_y} \left[ \delta\hat{\phi}_{k_x, k_y}(\theta(\mathbf{R}), t) J_0 \left( \frac{u_0 |\mathbf{k}_{\perp}|}{\Omega} \right) \right] e^{ik_x x(\mathbf{R})} e^{ik_y y(\mathbf{R})} + \mathcal{O}(\rho_* \delta\phi) \quad (\text{D.118})$$

and so we find that in Fourier space, gyroaveraging may be performed analytically and amounts to multiplying by a Bessel function.

For gyroaveraging  $h_s$  at constant particle position  $\mathbf{r}$ , as appears in equation 3.106, through analogous analysis one can show

$$\langle h_s(\mathbf{r} - \boldsymbol{\rho}_0, t) \rangle_{\mathbf{r}} = \sum_{k_x, k_y} \left[ \hat{h}_{s, k_x, k_y}(\theta(\mathbf{r}), t) J_0 \left( \frac{u_0 |\mathbf{k}_{\perp}|}{\Omega_s} \right) \right] e^{ik_x x(\mathbf{r})} e^{ik_y y(\mathbf{r})} + \mathcal{O}(\rho_* \delta f_s). \quad (\text{D.119})$$

# Appendix E

## Developing SAT3

### E.1 Parseval's theorem

Parseval's theorem relates the average of an absolute squared function in real space to the sum of the absolute squares of its Fourier amplitudes. Consider the Fourier representation of a function  $f(x)$  over a finite domain  $\{x_0, x_0 + L\}$

$$f(x) = \sum_{k_x} \hat{f}_{k_x} e^{ik_x x} \quad (\text{E.1})$$

where  $k_x = 2\pi n_x/L$  for all integers  $n_x$ . Multiplying both sides by  $f^*(x)$  and averaging over the domain leaves only the  $k_x = k'_x$  terms, producing the result.

$$\begin{aligned} \frac{1}{L} \int_{x_0}^{x_0+L} |f(x)|^2 dx &= \sum_{k_x} \sum_{k'_x} \hat{f}_{k_x} \hat{f}_{k'_x}^* \frac{1}{L} \int_{x_0}^{x_0+L} e^{i(k_x - k'_x)x} dx \\ &= \sum_{k_x} \left| \hat{f}_{k_x} \right|^2 \end{aligned} \quad (\text{E.2})$$

### E.2 Grid-resolution dependence

The flux  $Q_s$  in equation 4.1 is a physical quantity, resulting in part from a spatial average over the  $x$  and  $y$  domains. It should therefore be independent of the binormal wavenumber grid resolution  $\Delta k_y$ , once past a certain value required for convergence. It then follows that the flux components  $Q_{s,k_y}$  must depend on the grid resolution, as if one doubles the resolution, one doubles the number of terms in the sum while maintaining a constant total.

Multiplying and dividing the right-hand side of equation 4.1 by  $\Delta k_y$  and approximating the

summation as an integral gives

$$Q_s = \sum_{k_y > 0} \frac{Q_{s,k_y}}{\Delta k_y} \Delta k_y \approx \int_0^\infty \frac{Q_{s,k_y}}{\Delta k_y} dk_y \quad (\text{E.3})$$

and thus because the area under  $Q_{s,k_y}/\Delta k_y$  is independent of the grid resolution, so too must  $Q_{s,k_y}/\Delta k_y$  be itself. Analogously for the potentials, the volume and time average of the squared potentials in real space must be independent of  $\Delta k_y$ . Using Parseval's theorem,

$$\langle |\delta\phi|^2 \rangle_{x,y,\theta,t} = \sum_{k_y} \langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t} \quad (\text{E.4})$$

it follows that

$$\langle |\delta\phi|^2 \rangle_{x,y,\theta,t} \approx \int_{-\infty}^{\infty} \frac{\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t}}{\Delta k_y} dk_y \quad (\text{E.5})$$

and therefore  $\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t} / \Delta k_y$  must also be independent of  $\Delta k_y$ . Now using equation 3.215, one finds

$$\frac{\langle |\delta\hat{\phi}_{k_y}|^2 \rangle_{x,\theta,t}}{\Delta k_y} = \sum_{k_x} \frac{\langle |\delta\hat{\phi}_{k_x,k_y}|^2 \rangle_{\theta,t}}{\Delta k_y \Delta k_x} \Delta k_x \approx \int_{-\infty}^{\infty} \frac{\langle |\delta\hat{\phi}_{k_x,k_y}|^2 \rangle_{\theta,t}}{\Delta k_y \Delta k_x} dk_x \quad (\text{E.6})$$

resulting in  $\langle |\delta\hat{\phi}_{k_x,k_y}|^2 \rangle_{\theta,t} / \Delta k_y \Delta k_x$  being the grid-independent version of the 2D potentials.

### E.3 Database convergence study

In preparing the nonlinear gyrokinetic database convergence studies were performed, looking at the grid resolutions required in the radial and binormal directions for the convergence of fluxes. This was done by keeping the maximum values of  $k_y \rho_{\text{unit}}$  and  $k_x \rho_{\text{unit}}$  approximately constant and varying  $N_y$  and  $N_x$ . This varies the minimum values of the wavenumbers in the simulation, which correspond to the size of the domain in physical space, for example  $(k_x \rho_{\text{unit}})_{\text{min}} = 2\pi \rho_{\text{unit}} / L_x$ .

From a GA-std case in D baseline, then the integer parameter  $S$  was used to change the resolutions, via

$$(k_y \rho_{\text{unit}})_{\text{min}} = \frac{1.0}{8S - 1}, \quad N_x = 64S, \quad N_y = 8S \quad (\text{E.7})$$

in which  $S$  was varied between 1 and 5. By keeping the box size integer constant at  $w = 6$  (equation 3.188), this variation of  $S$  keeps the same  $(k_y \rho_{\text{unit}})_{\text{max}} = 1.0$ ,  $(k_x \rho_{\text{unit}})_{\text{max}} = (N_x - 2)(k_x \rho_{\text{unit}})_{\text{min}} / 2 = (\pi/6)(64S - 2) / (8S - 1) \approx 4.3$  while changing the minimum wavenumbers,  $(k_y \rho_{\text{unit}})_{\text{min}} = 1 / (8S - 1)$  and  $(k_x \rho_{\text{unit}})_{\text{min}} = (2\pi/6) / (8S - 1)$ .

The ion heat flux components normalised to  $\Delta k_y$  for this exercise are shown in E.1a. By the arguments of appendix E.2, this function should be unchanged by a change in resolution once converged. This can be seen to be the case for  $S \geq 3$ , however the results of  $S = 1$  and  $S = 2$  indicate a lack of convergence. Note that the  $S = 2$  result is indicative of resolutions used in previous studies [64]. How the total fluxes vary with  $S$  is then shown in E.1b, from which it can be seen the difference between the flux for  $S = 2$  to  $S = 5$  is an increase of  $\sim 25\%$ .

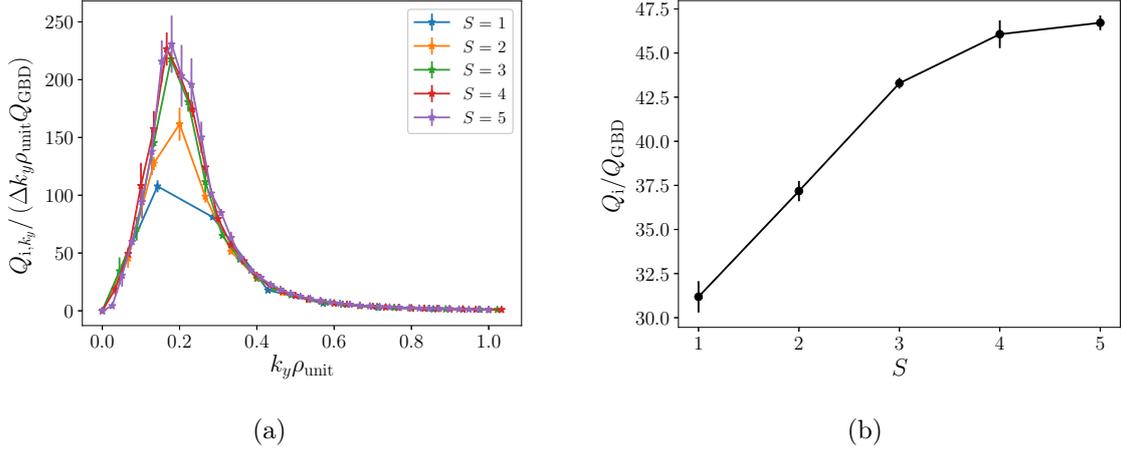


Figure E.1: (a) Normalised ion heat flux components against  $k_y \rho_{\text{unit}}$  for the GA-std case in D, varying the resolution parameter  $S$ . (b) The total ion heat fluxes for the cases in (a), plotted against  $S$ .

While these results show that the results for  $S < 3$  are not converged, this does not inform us as to whether this originates from the binormal direction, the radial direction, or a combination of the two. To determine this, then keeping  $(k_y \rho_{\text{unit}})_{\text{max}} = 1.0$  and  $N_y = 40$ , the parameters  $N_x$  and  $w$  were changed so as to vary the radial resolution for an approximately constant maximum radial wavenumber and a constant binormal domain. The pairs of  $\{w, N_x\}$  used were  $(\{2, 108\}, \{3, 160\}, \{4, 216\}, \{6, 320\})$ . If the radial domain is underresolved for the  $S = 1, 2$  cases, then this exercise should show a decrease in flux for decreasing  $w$ . However, as seen in figure E.2a, this is not the case. Varying  $w$  from 6 to 2 has minimal impact on the fluxes, indicating that it is the binormal domain that is underresolved in the cases of  $S = 1, 2$ . This is further supported by the results of figure E.2b, which shows a similar study for the  $a/L_n = 3.0$  case in H. Again for a constant binormal domain with  $N_y = 24$ , then  $N_x$  and  $w$  were varied via  $N_x = 64S_{\text{rad}}$  and  $w = 2S_{\text{rad}}$ . Varying  $S_{\text{rad}}$  between 1 and 9 produced no significant deviation in the fluxes, indicating again that the convergence is mainly associated with the binormal domain.

Based on these observations the resolutions of  $N_y = 40$  and  $N_x = 224$ ,  $w = 4$  were used for

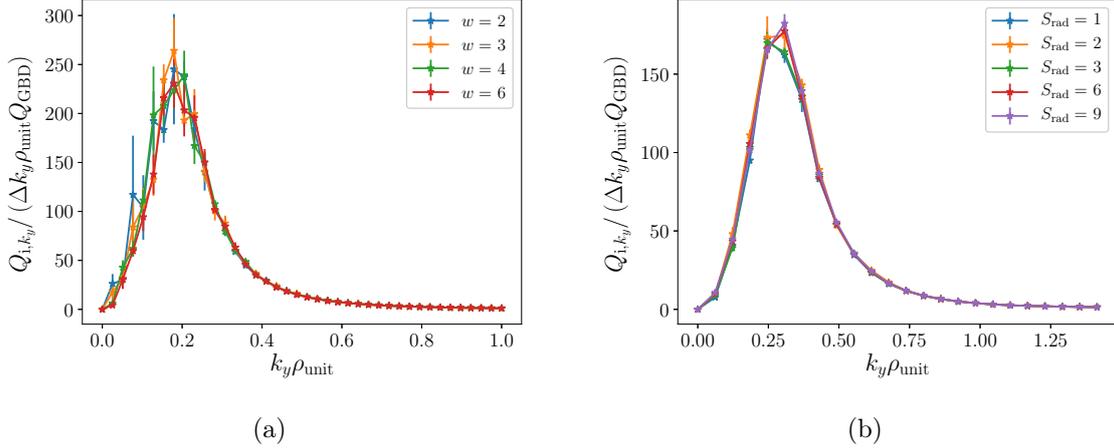


Figure E.2: (a) Normalised ion heat flux for the GA-std case in D against  $k_y \rho_{\text{unit}}$  for different radial wavenumber resolutions. Note that the flux components are essentially unchanged between the simulations, indicating no issue of radial convergence. (b) A similar study for the  $a/L_n = 3.0$  case in H, showing no appreciable change in the flux components across a ninefold change in radial resolution.

the database. This maintains a radial resolution similar to those of previous studies, but with a greater binormal resolution.

## E.4 2D spectrum moment integrals

Here the first three moments of the 2D potential spectrum (equations 5.1, 5.4 and 5.5) are integrated analytically to obtain expressions for the coefficients  $C_1$ ,  $C_2$  and  $K$  in equation 5.3. Starting with the zeroth moment, defined by equation 5.1, then by approximating the summation as an integral and inserting equation 5.3 for the 2D potential spectrum, one obtains

$$\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t} = \frac{1}{\Delta k_x} \int_{-\infty}^{\infty} \frac{\left\langle \left| \delta \hat{\phi}_{k_x=K,k_y} \right|^2 \right\rangle_{\theta,t}}{1 + C_1 (k_x - K)^2 + C_2 (k_x - K)^4} dk_x. \quad (\text{E.8})$$

Taking out a factor of  $C_2$  from the denominator, the integral may be written

$$\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t} = D \int_{-\infty}^{\infty} \frac{1}{E + Fu^2 + u^4} du \quad (\text{E.9})$$

where  $D = \left\langle \left| \delta \hat{\phi}_{k_x=K,k_y} \right|^2 \right\rangle_{\theta,t} / (C_2 \Delta k_x)$ ,  $E = 1/C_2$ ,  $F = C_1/C_2$  and  $u = k_x - K$ . The denominator can then be factorised via  $E + Fu^2 + u^4 = (u^2 + G)(u^2 + H)$ , where  $G = \frac{1}{2} (F + \sqrt{F^2 - 4E})$ ,  $H = \frac{1}{2} (F - \sqrt{F^2 - 4E})$  and  $F = G + H$ ,  $E = GH$ . Note that for the integral to have no singularities one requires  $G, H > 0$ , and hence  $E, F > 0$ . This factorised

form can then be separated using partial fractions as

$$\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t} = \frac{D}{G-H} \int_{-\infty}^{\infty} \left[ \frac{1}{u^2+H} - \frac{1}{u^2+G} \right] du. \quad (\text{E.10})$$

Using the substitution  $u = \sqrt{H} \tan(v)$  and  $u = \sqrt{G} \tan(v)$  respectively for the two integrals, equation E.10 evaluates to

$$\begin{aligned} \left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t} &= \frac{D}{G-H} \left[ \frac{1}{\sqrt{H}} \arctan \left( \frac{u}{\sqrt{H}} \right) - \frac{1}{\sqrt{G}} \arctan \left( \frac{u}{\sqrt{G}} \right) \right]_{-\infty}^{\infty} \\ &= \frac{D\pi}{\sqrt{GH} (\sqrt{G} + \sqrt{H})}. \end{aligned} \quad (\text{E.11})$$

Squaring this equation, using the relations  $GH = 1/C_2$ ,  $G + H = C_1/C_2$  and re-arranging, one finds

$$C_1 + 2\sqrt{C_2} = \left( \frac{\pi \left\langle \left| \delta \hat{\phi}_{k_x=K, k_y} \right|^2 \right\rangle_{\theta,t}}{\Delta k_x \left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \right)^2. \quad (\text{E.12})$$

Turning now to the first moment, equation 5.4, this can be found via

$$\begin{aligned} \langle k_x \rangle &= \frac{\sum_{k_x} (u + K) \left\langle \left| \delta \hat{\phi}_{k_x, k_y} \right|^2 \right\rangle_{\theta,t}}{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \\ &= \frac{\sum_{k_x} u \left\langle \left| \delta \hat{\phi}_{k_x, k_y} \right|^2 \right\rangle_{\theta,t}}{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} + K \\ &= \frac{D}{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \left[ \int_{-\infty}^{\infty} \frac{u}{E + Fu^2 + u^4} du \right] + K \\ &= K \end{aligned} \quad (\text{E.13})$$

as the integral of an odd function over a symmetric boundary is zero.

The second moment, equation 5.5, can be written

$$\sigma_{k_y}^2 = \frac{D}{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \int_{-\infty}^{\infty} \frac{u^2}{E + Fu^2 + u^4} du \quad (\text{E.14})$$

which, following an analogous method of solution as the zeroth moment, evaluates to

$$\begin{aligned}
\sigma_{k_y}^2 &= \frac{1}{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \frac{D}{G-H} \int_{-\infty}^{\infty} \left[ \frac{G}{u^2+G} - \frac{H}{u^2+H} \right] du \\
&= \frac{1}{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \frac{D}{G-H} \left[ \sqrt{G} \arctan \left( \frac{u}{\sqrt{G}} \right) - \sqrt{H} \arctan \left( \frac{u}{\sqrt{H}} \right) \right]_{-\infty}^{\infty} \\
&= \frac{1}{\left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \frac{D\pi}{\sqrt{G} + \sqrt{H}}.
\end{aligned} \tag{E.15}$$

Squaring this result and expressing it in terms of  $C_1$  and  $C_2$  gives

$$C_2 \left( C_1 + 2\sqrt{C_2} \right) = \left( \frac{\pi \left\langle \left| \delta \hat{\phi}_{k_x=\langle k_x \rangle, k_y} \right|^2 \right\rangle_{\theta,t}}{\Delta k_x \left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \right)^2 \frac{1}{\sigma_{k_y}^4} \tag{E.16}$$

which when combined with equation E.12, produces the desired solutions

$$C_2 = \frac{1}{\sigma_{k_y}^4}, \quad C_1 = \left[ \left( \frac{\pi \left\langle \left| \delta \hat{\phi}_{k_x=\langle k_x \rangle, k_y} \right|^2 \right\rangle_{\theta,t} \sigma_{k_y}}{\Delta k_x \left\langle \left| \delta \hat{\phi}_{k_y} \right|^2 \right\rangle_{x,\theta,t}} \right)^2 - 2 \right] \frac{1}{\sigma_{k_y}^2}. \tag{E.17}$$

The results of this section have been derived assuming integration over the entire  $k_x$  domain, which is inaccessible for turbulence simulations. For some finite domain  $\{-k_{x,\text{lim}}, k_{x,\text{lim}}\}$ , these results continue to hold provided that  $k_{x,\text{lim}} \gg |\langle k_x \rangle| + \sigma_{k_y}$ .

# Bibliography

- [1] L. Lucas and M. Unterweger, *Journal of Research of the National Institute of Standards and Technology* **105**, 541 (2000).
- [2] J. D. Huba, *NRL Plasma Formulary* (Naval Research Laboratory, 2013).
- [3] *Nuclear fusion cross sections and reactivities*.
- [4] *ENDF: Evaluated Nuclear Data File*.
- [5] R. Hagemann, G. Nief, and E. Roth, *Tellus* **22**, 712 (1970).
- [6] J. P. Freidberg, *Plasma Physics and Fusion Energy* (Cambridge University Press, 2008).
- [7] J. D. Lawson, *Proceedings of the Physical Society. Section B* **70**, 6 (1957).
- [8] J. Wesson and D. J. Campbell, *Tokamaks* (OUP Oxford, 2011).
- [9] T. J. M. Boyd and J. J. Sanderson, *The Physics of Plasmas* (Cambridge University Press, 2003).
- [10] S. I. Braginskii, *Reviews of Plasma Physics* **1**, 205 (1965).
- [11] R. D. Hazeltine and J. D. Meiss, *Plasma Confinement* (Courier Corporation, 2013).
- [12] A. G. Peeters, *The physics of fusion power* (2008).
- [13] S. Li, H. Jiang, Z. Ren, et al., *Abstract and Applied Analysis* **2014**, 1 (2014).
- [14] F. J. Casson, PhD in physics, University of Warwick (2011).
- [15] M. Keilhacker, A. Gibson, C. Gormezano, et al., *Nuclear Fusion* **39**, 209 (1999).
- [16] G. P. Thomson and M. Blackman, *Improvements in or relating to gas discharge apparatus for producing thermonuclear reactions* (1959).
- [17] R. Aymar, V. Chuyanov, M. Huguet, et al., *Nuclear Fusion* **41**, 1301 (2001).

- [18] ITER Physics Basis Expert Group on Confinement and Transport et al., **39**, 2175 (1999).
- [19] J. Mailloux, N. Abid, K. Abraham, et al., *Nuclear Fusion* **62**, 042026 (2022).
- [20] TFR Group and A. Truc, *Nuclear Fusion* **26**, 1303 (1986).
- [21] A. J. Wootton, B. A. Carreras, H. Matsumoto, et al., *Physics of Fluids B: Plasma Physics* **2**, 2879 (1990).
- [22] R. J. Fonck, G. Cosby, R. D. Durst, et al., *Physical Review Letters* **70**, 3736 (1993).
- [23] E. Mazzucato, *Physical Review Letters* **36**, 792 (1976).
- [24] W. Horton, *Reviews of Modern Physics* **71**, 735 (1999).
- [25] J. Weiland, *Collective Modes in Inhomogeneous Plasmas: Kinetic and Advanced Fluid Theory* (CRC Press, 1999).
- [26] J. E. Rice, J. Citrin, N. M. Cao, et al., *Nuclear Fusion* **60**, 105001 (2020).
- [27] S. M. Kaye, M. Greenwald, U. Stroth, et al., *Nuclear Fusion* **37**, 1303 (1997).
- [28] F. Wagner, G. Becker, K. Behringer, et al., *Physical Review Letters* **49**, 1408 (1982).
- [29] J. W. Connor and H. R. Wilson, *Plasma Physics and Controlled Fusion* **42**, R1 (1999).
- [30] H. Zohm, *Plasma Physics and Controlled Fusion* **38**, 105 (1996).
- [31] I. G. Abel, G. G. Plunk, E. Wang, et al., *Reports on Progress in Physics* **76**, 116201 (2013).
- [32] G. Cenacchi and A. Taroni, *Jetto: a free boundary plasma transport code* (1988).
- [33] G. V. Pereverzev and P. N. Yushmanov, *ASTRA - Automated System for TRansport Analysis* (2002).
- [34] A. J. Brizard and T. S. Hahm, *Reviews of Modern Physics* **79**, 421 (2007).
- [35] A. G. Peeters, Y. Camenen, F. J. Casson, et al., *Computer Physics Communications* **180**, 2650 (2009).
- [36] X. Garbet, Y. Idomura, L. Villard, et al., *Nuclear Fusion* **50**, 043002 (2010).
- [37] J. Candy and E. A. Belli, *Gyro Technical Guide* (2011).

- [38] J. Candy, E. A. Belli, and R. V. Bravenec, *Journal of Computational Physics* **324**, 73 (2016).
- [39] T. Rhodes, C. Holland, S. Smith, et al., *Nuclear Fusion* **51**, 063022 (2011).
- [40] C. Holland, L. Schmitz, T. L. Rhodes, et al., *Physics of Plasmas* **18**, 056113 (2011).
- [41] D. Told, F. Jenko, T. Görler, et al., *Physics of Plasmas* **20**, 122312 (2013).
- [42] J. Citrin, F. Jenko, P. Mantica, et al., *Nuclear Fusion* **54**, 023008 (2014).
- [43] N. Bonanomi, P. Mantica, G. Szepesi, et al., *Nuclear Fusion* **55**, 113016 (2015).
- [44] A. J. Creely, N. T. Howard, P. Rodriguez-Fernandez, et al., *Physics of Plasmas* **24**, 056104 (2017).
- [45] S. J. Freethy, T. Görler, A. J. Creely, et al., *Physics of Plasmas* **25**, 055903 (2018).
- [46] J. Citrin, S. Maeyama, C. Angioni, et al., *Nuclear Fusion* **62**, 086025 (2022).
- [47] M. Barnes, PhD in physics, University of Maryland (2009).
- [48] J. Candy, R. Waltz, and M. Fahey, in *50th Annual Meeting of the Division of Plasma Physics* (2008).
- [49] C. Bourdelle, J. Citrin, B. Baiocchi, et al., *Plasma Physics and Controlled Fusion* **58**, 014036 (2015).
- [50] G. M. Staebler, J. E. Kinsey, and R. E. Waltz, *Physics of Plasmas* **12**, 102508 (2005).
- [51] A. Casati, C. Bourdelle, X. Garbet, et al., *Nuclear Fusion* **49**, 085012 (2009).
- [52] G. M. Staebler, E. A. Belli, J. Candy, et al., *Nuclear Fusion* **61**, 116007 (2021).
- [53] J. Citrin, C. Bourdelle, F. J. Casson, et al., *Plasma Physics and Controlled Fusion* **59**, 124005 (2017).
- [54] J. Garcia, R. Dumont, J. Joly, et al., *Nuclear Fusion* **59**, 086047 (2019).
- [55] A. Ho, J. Citrin, F. Auriemma, et al., *Nuclear Fusion* **59**, 056007 (2019).
- [56] F. Casson, H. Patten, C. Bourdelle, et al., *Nuclear Fusion* **60**, 066029 (2020).
- [57] M. Marin, J. Citrin, L. Garzotti, et al., *Nuclear Fusion* **61**, 036042 (2021).

- [58] C. C. Petty, T. C. Luce, R. I. Pinsker, et al., *Physical Review Letters* **74**, 1763 (1995).
- [59] M. Bessenrodt-Weberpals, F. Wagner, O. Gehre, et al., *Nuclear Fusion* **33**, 1205 (1993).
- [60] P. A. Schneider, A. Bustos, P. Hennequin, et al., *Nuclear Fusion* **57**, 066003 (2017).
- [61] C. F. Maggi, H. Weisen, J. C. Hillesheim, et al., *Plasma Physics and Controlled Fusion* **60**, 014045 (2017).
- [62] H. Weisen, C. F. Maggi, M. Oberparleiter, et al., *Journal of Plasma Physics* **86**, 905860501 (2020).
- [63] G. Manfredi and M. Ottaviani, *Physical Review Letters* **79**, 4190 (1997).
- [64] E. A. Belli, J. Candy, and R. E. Waltz, *Physics of Plasmas* **26**, 082305 (2019).
- [65] J. Garcia, T. Görler, F. Jenko, et al., *Nuclear Fusion* **57**, 014007 (2016).
- [66] M. Nakata, M. Nunami, H. Sugama, et al., *Physical Review Letters* **118**, 165002 (2017).
- [67] I. Pusztai, J. Candy, and P. Gohil, *Physics of Plasmas* **18**, 122501 (2011).
- [68] J. Garcia, F. J. Casson, A. Bañón Navarro, et al., *Plasma Physics and Controlled Fusion* **64**, 054001 (2022).
- [69] L. Umlauf and H. Burchard, *Marine Turbulence* (2022).
- [70] A. C. C. Sips et al., *Plasma Physics and Controlled Fusion* **47**, A19 (2005).
- [71] H. Sugama and W. Horton, *Physics of Plasmas* **5**, 2560 (1998).
- [72] R. G. Littlejohn, *The Physics of Fluids* **24**, 1730 (1981).
- [73] A. Brizard, *Journal of Plasma Physics* **41**, 541 (1989).
- [74] P. J. Catto, *Plasma Physics* **20**, 719 (1978).
- [75] J. Candy and R. E. Waltz, *Journal of Computational Physics* **186**, 545 (2003).
- [76] F. Jenko, W. Dorland, M. Kotschenreuther, et al., *Physics of Plasmas* **7**, 1904 (2000).
- [77] E. A. Frieman and L. Chen, *The Physics of Fluids* **25**, 502 (1982).
- [78] S. Cowley, *From ITER Time-scales to Gyro-Kinetics*.

- [79] G. G. Howes, S. C. Cowley, W. Dorland, et al., *The Astrophysical Journal* **651**, 590 (2006).
- [80] I. G. Abel, M. Barnes, S. C. Cowley, et al., *Physics of Plasmas* **15**, 122509 (2008).
- [81] H. Sugama, T.-H. Watanabe, and M. Nunami, *Physics of Plasmas* **16**, 112503 (2009).
- [82] B. Li and D. R. Ernst, *Physical Review Letters* **106**, 195002 (2011).
- [83] F. L. Hinton and S. K. Wong, *The Physics of Fluids* **28**, 3082 (1985).
- [84] M. Artun and W. M. Tang, *Physics of Plasmas* **1**, 2682 (1994).
- [85] W. D. D'haeseleer, *Flux Coordinates and Magnetic Field Structure: A Guide to a Fundamental Tool of Plasma Theory* (Springer-Verlag, 1991).
- [86] M. A. Beer, S. C. Cowley, and G. W. Hammett, *Physics of Plasmas* **2**, 2687 (1995).
- [87] S. Hamada, *Nuclear Fusion* **2**, 23 (1962).
- [88] J. M. Greene and J. L. Johnson, *The Physics of Fluids* **5**, 510 (1962).
- [89] C. Mercier and N. Luc, Tech. Rep. EUR-5127e 140, Commission of the European Communities (1974).
- [90] J. Candy, *Plasma Physics and Controlled Fusion* **51**, 105009 (2009).
- [91] J. Ball and F. I. Parra, *Plasma Physics and Controlled Fusion* **58**, 055016 (2016).
- [92] R. L. Miller, M. S. Chu, J. M. Greene, et al., *Physics of Plasmas* **5**, 973 (1998).
- [93] A. D. Turnbull, Y. R. Lin-Liu, R. L. Miller, et al., *Physics of Plasmas* **6**, 1113 (1999).
- [94] M. A. Beer, PhD in physics, Princeton University (1995).
- [95] R. E. Waltz and R. L. Miller, *Physics of Plasmas* **6**, 4265 (1999).
- [96] W. W. Lee, *The Physics of Fluids* **26**, 556 (1983).
- [97] A. A. Schekochihin, S. C. Cowley, W. Dorland, et al., *The Astrophysical Journal Supplement Series* **182**, 310 (2009).
- [98] M. F. Hoyaux, *American Journal of Physics* **35**, 232 (1967).
- [99] R. E. Waltz, G. M. Staebler, W. Dorland, et al., *Physics of Plasmas* **4**, 2482 (1997).

- [100] R. E. Waltz, G. D. Kerbel, and J. Milovich, *Physics of Plasmas* **1**, 2229 (1994).
- [101] Z. Lin, T. S. Hahm, W. W. Lee, et al., *Science* **281**, 1835 (1998).
- [102] R. E. Waltz, A. Casati, and G. M. Staebler, *Physics of Plasmas* **16**, 072303 (2009).
- [103] A. Casati, PhD in physics, University of Provence (2012).
- [104] M. A. Beer and G. W. Hammett, *Physics of Plasmas* **3**, 4046 (1996).
- [105] P. B. Snyder, PhD in physics, Princeton University (1999).
- [106] F. Merz and F. Jenko, *Nuclear Fusion* **50**, 054005 (2010).
- [107] D. R. Hatch, P. W. Terry, F. Jenko, et al., *Physics of Plasmas* **18**, 055706 (2011).
- [108] K. H. Burrell, *Physics of Plasmas* **4**, 1499 (1997).
- [109] G. V. Oost, J. Ad mek, V. Antoni, et al., *Plasma Physics and Controlled Fusion* **45**, 621 (2003).
- [110] R. E. Waltz, G. M. Staebler, J. Candy, et al., *Physics of Plasmas* **14**, 122507 (2007).
- [111] R. Buchholz, PhD in physics, University of Bayreuth (2015).
- [112] F. J. Casson, A. G. Peeters, Y. Camenen, et al., *Physics of Plasmas* **16**, 092303 (2009).
- [113] T. Dannert and F. Jenko, *Physics of Plasmas* **12**, 072309 (2005).
- [114] C. Angioni, A. G. Peeters, F. Jenko, et al., *Physics of Plasmas* **12**, 112310 (2005).
- [115] F. Casson, R. McDermott, C. Angioni, et al., *Nuclear Fusion* **53**, 063026 (2013).
- [116] B. Patel, PhD in physics, University of York (2021).
- [117] R. E. Waltz, R. R. Dominguez, and G. W. Hammett, *Physics of Fluids B: Plasma Physics* **4**, 3138 (1992).
- [118] G. M. Staebler, J. Candy, N. T. Howard, et al., *Physics of Plasmas* **23**, 062518 (2016).
- [119] N. Howard, C. Holland, A. White, et al., *Nuclear Fusion* **56**, 014004 (2016).
- [120] W. W. Lee, J. A. Krommes, C. R. Oberman, et al., *The Physics of Fluids* **27**, 2652 (1984).

- [121] M. Barnes, F. I. Parra, E. G. Highcock, et al., *Physical Review Letters* **106**, 175004 (2011).
- [122] G. M. Staebler, R. E. Waltz, J. Candy, et al., *Physical Review Letters* **110**, 055003 (2013).
- [123] G. Staebler, J. Candy, R. Waltz, et al., *Nuclear Fusion* **53**, 113017 (2013).
- [124] R. E. Waltz, R. L. Dewar, and X. Garbet, *Physics of Plasmas* **5**, 1784 (1998).
- [125] C. Bourdelle, PhD in physics, Joseph Fourier University (2000).
- [126] C. Bourdelle, X. Garbet, G. T. Hoang, et al., *Nuclear Fusion* **42**, 892 (2002).
- [127] J. W. Connor, R. J. Hastie, and T. J. Martin, *Nuclear Fusion* **23**, 1702 (1983).
- [128] C. Stephens, X. Garbet, J. Citrin, et al., *Journal of Plasma Physics* **87**, 905870409 (2021).
- [129] J. Citrin, *QuaLiKiz Saturation Rule* (2019).
- [130] J. Citrin, C. Bourdelle, P. Cottier, et al., *Physics of Plasmas* **19**, 062305 (2012).
- [131] V. D. Shafranov, *Soviet Atomic Energy* **13**, 1149 (1963).
- [132] C. J. Ajay, S. Brunner, B. McMillan, et al., *Journal of Plasma Physics* **86**, 905860504 (2020).
- [133] A. Bustos, A. Bañón Navarro, T. Görler, et al., *Physics of Plasmas* **22**, 012305 (2015).
- [134] A. G. Peeters, C. Angioni, A. Bortolon, et al., *Nuclear Fusion* **51**, 094027 (2011).
- [135] P. Cottier, C. Bourdelle, Y. Camenen, et al., *Plasma Physics and Controlled Fusion* **56**, 015011 (2014).
- [136] C. Holland, N. Howard, and B. Grierson, *Nuclear Fusion* **57**, 066043 (2017).
- [137] H. G. Dudding, F. J. Casson, D. Dickinson, et al., *Nuclear Fusion* **62**, 096005 (2022).
- [138] J. Figueiredo, A. Murari, C. Perez Von Thun, et al., *Review of Scientific Instruments* **89**, 10K119 (2018).
- [139] G. F. Matthews, P. Edwards, T. Hirai, et al., *Physica Scripta* **2007**, 137 (2007).
- [140] W. A. Houlberg, K. C. Shaing, S. P. Hirshman, et al., *Physics of Plasmas* **4**, 3230 (1997).

- [141] C. D. Challis, J. G. Cordey, H. Hamnén, et al., *Nuclear Fusion* **29**, 563 (1989).
- [142] R. Pasqualotto, P. Nielsen, C. Gowers, et al., *Review of Scientific Instruments* **75**, 3891 (2004).
- [143] L. L. Lao, H. S. John, R. D. Stambaugh, et al., *Nuclear Fusion* **25**, 1611 (1985).
- [144] C. L. Rettig, T. L. Rhodes, J. N. Leboeuf, et al., *Physics of Plasmas* **8**, 2232 (2001).
- [145] E. Delabie, *To be submitted*.
- [146] R. Hawryluk, in *Physics of Plasmas Close to Thermonuclear Conditions* (1981).
- [147] E. Delabie, M. M. F. Nave, M. Baruzzo, et al., in *44th EPS Conference on Plasma Physics* (2017).
- [148] E. Delabie, M. M. F. Nave, P. Rodriguez-Fernandez, et al., **2021**, CO07.003 (2021).
- [149] R. J. Hastie, *Astrophysics and Space Science* **256**, 177 (1997).
- [150] C. Maggi, H. Weisen, F. Casson, et al., *Nuclear Fusion* **59**, 076028 (2019).
- [151] A. M. Dimits, G. Bateman, M. A. Beer, et al., *Physics of Plasmas* **7**, 969 (2000).
- [152] X. Garbet, P. Mantica, F. Ryter, et al., *Plasma Physics and Controlled Fusion* **46**, 1351 (2004).
- [153] E. A. Belli, J. Candy, and R. E. Waltz, *Physical Review Letters* **125**, 015001 (2020).
- [154] D. R. Ernst, B. Coppi, S. D. Scott, et al., *Physical Review Letters* **81**, 2454 (1998).
- [155] ITER Organisation, *ITER Research Plan within the Staged Approach (Level III – Provisional Version)* (2018).
- [156] N. Kumar, Y. Camenen, S. Benkadda, et al., *Nuclear Fusion* **61**, 036005 (2021).
- [157] T. Rafiq, S. Kaye, W. Guttenfelder, et al., *Physics of Plasmas* **28**, 022504 (2021).
- [158] F. L. Hinton and R. D. Hazeltine, *Reviews of Modern Physics* **48**, 239 (1976).
- [159] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables* (Courier Corporation, 1965).