

Particle acceleration in Kerr black hole magnetospheres

Benjamin Crinquand

▶ **To cite this version:**

Benjamin Crinquand. Particle acceleration in Kerr black hole magnetospheres. Astrophysics [astro-ph]. Université Grenoble Alpes [2020-..], 2021. English. NNT : 2021GRALY045 . tel-03406333v2

HAL Id: tel-03406333

https://tel.archives-ouvertes.fr/tel-03406333v2

Submitted on 14 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : Astronomie et Astrophysique

Arrêté ministériel : 25 mai 2016

Présentée par

Benjamin CRINQUAND

Thèse dirigée par **Guillaume DUBUS, directeur de recherche, CNRS**, et
encadrée par **Benoît CERUTTI, chargé de recherche, CNRS**

préparée au sein de l'**Institut de Planétologie et d'Astrophysique de Grenoble (IPAG)**
dans **l'École Doctorale de Physique de Grenoble**

Accélération de particules dans les magnétosphères de trous noirs de Kerr

Thèse soutenue publiquement le **vendredi 9 juillet 2021**,
devant le jury composé de :

M., Frédéric, DAIGNE
Professeur à Sorbonne université, rapporteur.
M., Gilles, HENRI
Professeur à l'université Grenoble Alpes, président du jury.
M., Serguei, KOMISSAROV
Professeur à l'université de Leeds, rapporteur.
Mme, Monika, MOŚCIBRODZKA
Professeure assistante à l'université Radboud, examinatrice.
M., Frédéric, VINCENT
Chargé de recherche à l'Observatoire de Paris, examinateur.

UNIVERSITÉ GRENOBLE ALPES

DOCTORAL THESIS

Particle acceleration in Kerr black-hole magnetospheres

Author:
Benjamin CRINQUAND

Supervisors:
Benoît CERUTTI &
Guillaume DUBUS





"Je laisse Sisyphe au bas de la montagne ! On retrouve toujours son fardeau. Mais Sisyphe enseigne la fidélité supérieure qui nie les dieux et soulève les rochers. Lui aussi juge que tout est bien. Cet univers désormais sans maître ne lui paraît ni stérile, ni futile. Chacun des grains de cette pierre, chaque éclat minéral de cette montagne pleine de nuit, à lui seul forme un monde. La lutte en elle-même vers les sommets suffit à remplir un cœur d'homme. Il faut imaginer Sisyphe heureux."

"I leave Sisyphus at the foot of the mountain! One always finds one's burden again. But Sisyphus teaches the higher fidelity that negates the gods and raises rocks. He too concludes that all is well. This universe henceforth without a master seems to him neither sterile nor futile. Each atom of that stone, each mineral flake of that night filled mountain, in itself forms a world. The struggle itself toward the heights is enough to fill a man's heart. One must imagine Sisyphus happy."

ALBERT CAMUS, Le mythe de Sisyphe

Abstract

Black holes are the central engines of many high-energy astrophysical phenomena, such as active galactic nuclei (AGN). Black-hole activity often manifests itself through the launching of a relativistic jet filled with energetic particles. How jets form and accelerate particles remains poorly understood. Although black holes can only be observed indirectly through their impact on their environment, the situation has dramatically evolved in recent years due to formidable progress in astronomical instrumentation. The advent of ground-based Cherenkov telescopes proved that AGN could produce highly variable flares of non-thermal very high-energy radiation. Besides, the GRAVITY collaboration reported the detection of a flaring hot spot orbiting near the supermassive black hole Sgr A*, whereas radio interferometers have been able to track the jet emitted by the supermassive black hole M87* almost down to the event horizon. Then, in April 2019, the Event Horizon Telescope collaboration released the first image of the "shadow" of M87*. For the first time, the immediate vicinity of the black hole was spatially resolved.

To account for this activity, one of the most promising scenarios (the Blandford-Znajek mechanism) involves a rapidly spinning black hole immersed in a large-scale magnetic field. The dragging of spacetime and magnetic field lines by the black hole induces a potential drop between its poles, which can power an electromagnetic outflow and trigger particle acceleration. Thanks to the recent growth of computing power, it is now possible to test this scenario in great detail. Up to now, the magnetohydrodynamic approach has been mainly used to model black-hole magnetospheres. Such simulations have confirmed that the rotational energy of black holes can be electromagnetically extracted. However, they are unable to study particle acceleration by construction. Instead, we employ the "particle-in-cell" (PIC) approach to model a plasma from first principles. All the plasma processes involved in particle acceleration can be captured without physical approximations. This approach is relevant to collisionless plasmas, such as found near the supermassive black hole at the center of our Galaxy. PIC simulations have been successfully used to study particle acceleration in pulsar magnetospheres, but they had never been employed in this context.

In this thesis, I have used the general relativistic PIC code `Zeltron`, which is the first multidimensional kinetic code capable of evolving particles and fields in curved spacetime. I have upgraded `Zeltron` by including the possibility to track individual photons and implemented Monte-Carlo modules to treat radiative transfer accurately. I present global simulations of gaps in a black-hole magnetosphere, which prove that the Blandford-Znajek mechanism can indeed be activated self-consistently. I have pinpointed the location of the spark gaps and shown that pair creation occurs intermittently on scales smaller than the radius of the black hole. I also present simulations in a more realistic magnetic configuration, proving that magnetospheres can dissipate electromagnetic energy efficiently through magnetic reconnection. Besides, to bridge the gap between simulations and observations, I have coupled `Zeltron` with a ray-tracing code, allowing me to directly use the outcome of kinetic simulations to produce synthetic light curves and images.

Résumé

Les trous noirs sont à l'origine de nombreux phénomènes astrophysiques de haute énergie, tels que les noyaux actifs de galaxie (AGN). L'activité des trous noirs se manifeste souvent par le lancement d'un jet relativiste de particules énergétiques. Cependant, les mécanismes de formation des jets et d'accélération de particules restent mal compris. Si les trous noirs ne peuvent être observés qu'indirectement, à travers leur impact sur leur environnement, la situation a considérablement évolué ces dernières années grâce aux récents progrès instrumentaux. Des télescopes Tcherenkov ont détecté des éruptions de rayonnement non thermique de très haute énergie et variant sur de faibles échelles de temps en provenance d'AGN. Par ailleurs, la collaboration GRAVITY a rapporté la présence d'un point chaud en orbite près du trou noir supermassif Sgr A*, tandis que le jet émis par le trou noir supermassif M87* a pu être observé par interférométrie radio jusqu'à des échelles proches de l'horizon des événements du trou noir. Finalement, en avril 2019, la collaboration Event Horizon Telescope a publié la première image de « l'ombre » de M87*. Pour la première fois, le voisinage immédiat du trou noir a pu être résolu spatialement.

Pour modéliser cette activité, l'un des scénarios les plus prometteurs implique un trou noir en rotation rapide immergé dans un champ magnétique à grande échelle: c'est le mécanisme de Blandford-Znajek. L'entraînement de l'espace-temps et des lignes de champ magnétique par le trou noir induit une différence de potentiel entre ses pôles, pouvant alimenter un vent électromagnétique et déclencher l'accélération de particules. Les excellentes performances des supercalculateurs les plus récents permettent désormais de tester ce scénario dans les moindres détails. Jusqu'à présent, seule l'approche magnétohydrodynamique a été utilisée pour modéliser les magnétosphères de trous noirs. Ces simulations ont montré avec succès la viabilité du mécanisme d'extraction électromagnétique de l'énergie de rotation du trou noir. Cependant, elles demeurent incapables d'étudier l'accélération de particules énergétiques. Nous utilisons plutôt l'approche « particle-in-cell » (PIC), afin de modéliser un plasma à partir de principes fondamentaux. Tous les processus plasma impliqués dans l'accélération des particules peuvent être simulés sans aucune approximation physique. Une telle approche est justifiée lorsque le plasma est non collisionnel, comme aux environs du trou noir supermassif au centre de notre galaxie. Les simulations PIC ont été utilisées avec succès pour étudier l'accélération de particules dans les magnétosphères de pulsars, mais elles n'avaient jamais été employées dans ce contexte.

Dans cette thèse, j'ai utilisé le code PIC `Zeltron`, qui est le premier code cinétique multidimensionnel incluant complètement la relativité générale. J'ai amélioré `Zeltron` en y incluant la possibilité de suivre des photons individuels, et j'ai implémenté un module Monte-Carlo afin de traiter rigoureusement le transfert radiatif. Je présente des simulations globales de gaps dans une magnétosphère de trou noir, qui prouvent que le mécanisme de Blandford-Znajek peut effectivement être activé de manière auto-cohérente. J'ai localisé la position de ces gaps et montré que la création de paires se produit de manière intermittente, sur des échelles inférieures au rayon du trou noir. Je présente également des simulations dans une configuration magnétique plus réaliste, prouvant que l'énergie électromagnétique peut être dissipée très efficacement par reconnexion magnétique. Finalement, afin de combler le fossé entre les simulations et les observations, j'ai couplé `Zeltron` avec un code de tracé de rayons, ce qui me permet d'utiliser directement les résultats de simulations cinétiques pour produire des courbes de lumière et des images synthétiques.

Remerciements

Commençons par l'évidence. Beaucoup mériteraient d'être cités dans ces lignes : à l'échelle de ma vie, j'ai déjà reçu le concours de nombreuses personnes qui m'ont permis d'en arriver là. Pour des contraintes de place, et pour éviter aux lecteurs et lectrices de ces paragraphes une fastidieuse biographie, je me limiterai néanmoins aux évènements les plus récents et à des remerciements groupés. À celles et ceux dont le nom ne figure pas dans cette partie, je n'oublie pas ce que je leur dois.

Sans hésitation, je veux commencer par remercier mes deux directeurs de thèse, Benoît Cerutti et Guillaume Dubus, qui m'ont formé et guidé durant ces trois ans. C'est grâce à leur soutien et leur bienveillance que j'ai pu réussir cette thèse. Nous avons pu construire une relation basée sur la confiance. J'ignore si je peux légitimement me dire chercheur, mais grâce à eux j'aurai essayé, dans les meilleures conditions possibles. Par ailleurs, et ce n'est pas la moindre des choses, je leur suis très reconnaissant pour leur relecture rapide et consciencieuse de ce manuscrit. Avec quelques autres bonnes âmes qui ont accepté de se plonger dans l'ébauche de ce texte, et dont la contribution est dûment reconnue, ils m'ont aidé à débusquer nombre d'insidieuses coquilles et imprécisions.

Je remercie également chaudement les membres du jury, et en particulier les rapporteurs, pour avoir accepté de consacrer du temps et de l'énergie à évaluer mon travail. Nos échanges lors de la séance de questions furent très intéressants (et je ne manquerai plus de citer Breit et Wheeler !).

Je tiens particulièrement à remercier l'équipe SHERPAS, dont j'ai fait partie durant ces trois ans au sein de l'IPAG. L'atmosphère chaleureuse qui règne dans cette équipe et les bons moments partagés ont contribué à faciliter cette thèse. Même après cette année difficile où il aurait été facile de s'éloigner les uns des autres, le contact ne s'est jamais rompu. J'aurais seulement souhaité participer à plus de Sherpiades !

Je souhaite aussi remercier les autres doctorants de l'IPAG, compagnons de galère au quotidien. Leur flegme face à l'adversité, leurs bons conseils et leur excellente compagnie, même dans les moments difficiles, ont rendu ma vie grenobloise fort plaisante. Cette ambiance me manquera. Un autre groupe que je souhaite remercier particulièrement est le collectif JOLIDON. Un drôle de nom pour une drôle d'aventure, que nous avons finalement réussi à porter jusqu'au bout, ensemble. On peut dire que la route était droite, mais la pente était forte. Cependant, cela en valait la peine !

Je réalise que je dois également beaucoup à toutes ces personnes curieuses, amis proches ou brèves rencontres, qui m'ont posé la question fatidique : « Sur quoi porte ta thèse ? » (avec, de temps en temps, la variante « À quoi ça sert, ce que tu fais ? »). En tant que doctorant, savoir répondre à ces questions est indispensable, quel que soit le bagage scientifique de l'interlocuteur. Mais surtout, ces simples questions ont, de temps en temps, débouché sur de longs et fructueux débats. Voir des personnes a priori très éloignées de la physique et de l'astronomie montrer un sincère intérêt pour les problématiques que j'étudie est extrêmement gratifiant, et donne du sens à ce métier.

Finalement, à ma famille, mes amis et mes proches, pour avoir continué à partager avec moi leur part du monde pendant cette aventure grenobloise, merci. Et évidemment, à Émilie, qui a aussi dû supporter les mauvais côtés d'une thèse (les mauvaises humeurs quand rien ne marche, les week-ends de rédaction...), et qui est restée malgré tout, solaire, à mes côtés, merci.

Contents

Abstract iv

Résumé v

Remerciements vi

Foreword xi

Avant-propos xii

Conventions and notations xiii

Part I
Introduction: context and motivations

1 - A brief history of black holes 3

1.1 A theoretical concept . . . 3

1.2 The astrophysical age: stellar-mass black holes . . . 6

1.3 The astrophysical age: supermassive black holes . . . 10

1.4 Evidence for supermassive black holes in galactic nuclei . . . 15

2 - Multi-wavelength observations of supermassive black holes 18

2.1 Observations of jets . . . 18

2.2 The case of M87* . . . 24

2.3 The center of the Milky way: Sgr A* . . . 29

2.4 Very high-energy radiation from AGN . . . 33

2.5 Summary . . . 35

3 - State of the art 37

3.1 Magnetospheric gap models . . . 37

3.2 Numerical modeling . . . 40

3.3 Objectives of this thesis . . . 45

Part II
Foundations

4 - Kerr black holes 49

4.1 Spinning black holes . . . 49

4.2 Energy extraction . . . 55

4.A Appendix: Kerr-Schild coordinates . . . 57

5 - 3 + 1 electrodynamics 60

5.1 $3+1$ formulation of general relativity . . . 60

5.2 Maxwell's equations in general relativity . . . 65

5.3 $3+1$ reformulation of Maxwell's equations . . . 66

5.A Appendix: Useful notions on tensor analysis . . . 71

5.B Appendix: Covariant integration . . . 73

6 - Black-hole electrodynamics 75

6.1 Vacuum solutions . . . 75

6.2 The importance of plasma in the force-free approximation . . . 82

7 - The Blandford-Znajek mechanism 89

7.1 Michel's solution of a force-free pulsar magnetosphere . . . 89

7.2 The Blandford-Znajek solution . . . 91

7.A Appendix: Proof of the Znajek regularity condition . . . 99

8 - Radiative processes 101

8.1 Inverse Compton scattering . . . 101

8.2 Photon-photon annihilation . . . 108

8.3 Synchrotron radiation . . . 113

Part III
Numerical techniques

9 - Particle-in-cell simulations 121

9.1 Principles . . . 121
9.2 Particle motion . . . 123
9.3 Field evolution . . . 127
9.4 Parallelization . . . 132

10 - The `GRZeltron` upgrade 134

10.1 Particle solver . . . 134
10.2 Electromagnetic solver . . . 139
10.A Appendix: Derivation of the $3+1$ equations of motion . . . 143
10.B Appendix: Orthonormal tetrads in KS coordinates . . . 147

11 - Radiative transfer module 149

11.1 Principles . . . 149
11.2 Inverse Compton scattering . . . 150
11.3 Pair production . . . 154

Part IV
Results

12 - Simulations of pair discharges 159

12.1 Preliminaries . . . 159
12.2 Results . . . 163
12.3 Discussion . . . 171

13 - Dynamics of the current sheet 172

13.1 Numerical setup . . . 172
13.2 Results . . . 174
13.3 Magnetic field transport . . . 178
13.4 Toy model for magnetic flux decay . . . 180

13.A Appendix: Relativistic magnetic reconnection 182

14 - Synthetic observables 185

14.1 Null geodesics in Kerr spacetime 185

14.2 Light curves 192

14.3 Images 197

14.A Appendix: Black-hole shadow 205

Part V
Conclusion

15 - Conclusion and perspectives 211

Bibliography 215

Foreword

It has recently become especially rewarding and exciting to work in the field of relativistic astrophysics. In the last five years, several researchers were awarded a Nobel prize in physics for their studies on black holes. In 2017, Rainer Weiss, Barry C. Barish, and Kip S. Thorne received awards for their decisive contribution, within the LIGO collaboration, to the observation of gravitational waves. This detection has notably confirmed the existence of stellar-mass black holes. In 2020, Roger Penrose won the Nobel prize for his theoretical work on black holes as a prediction of the general theory of relativity, whereas Reinhard Genzel and Andrea Ghez were awarded for finding observational evidence of the presence of a supermassive black hole at the center of our galaxy. These awards illustrate how the scientific community has come to acknowledge the importance of black-hole physics.

Besides, the Event Horizon Telescope collaboration published in 2019 the first direct observation, at millimeter wavelengths, of “the shadow” of a black hole. This “photography” has aroused considerable interest in the world. The scientific communication surrounding this publication was impressive: the image was released during six simultaneous press conferences in America, Asia, and Europe. Thereafter, it was on the front page of most daily newspapers in the world. Aside from the scientific value of this observation, on which we will come back later, this media coverage highlights the public’s interest in black holes. These objects seem to hold a special place in the collective imagination and popular culture, as their properties defy one’s intuition.

Considering black holes as astrophysical objects has not always been self-evident. Even if the general theory of relativity has successfully passed many experimental tests, it was long believed that black holes were not a physical prediction of this theory. Starting from the first results of Albert Einstein and Karl Schwarzschild in 1915, it took several decades for the community to undergo a paradigm shift and to take black holes seriously despite their extreme properties. As Edwin Salpeter once said while the identification of the Cyg X-1 source was still controversial, in many cases the existence of a black hole must now be considered as “the most conservative hypothesis". It is precisely from this paradigm shift that the field of relativistic astrophysics was born.

In order to determine the nature of a light source in the sky, it is necessary to understand how this object interacts with its environment and to model accurately the radiation that it emits. It is one of the main themes of relativistic astrophysics to understand “what a black hole looks like”, depending on its environment. Far from being “cosmic vacuum cleaners" swallowing everything around them without leaving any observable trace, black holes power some of the most luminous sources in the observable universe. In particular, they are capable of producing plasma jets that can reach relativistic velocities and may extend over thousands of light-years. Future advances in astronomical instrumentation promise to unveil more details on their environment and to better constrain existing theoretical models. This thesis is part of an ongoing effort to interpret black hole observations by the means of numerical simulations. The observational and theoretical motivations of this thesis are outlined in Part I.

Avant-propos

Depuis quelques années, le domaine de l'astrophysique relativiste est en effeverscence. Au cours des cinq dernières années, plusieurs prix Nobel de physique ont été décernés à des chercheurs et chercheuses pour leurs travaux sur les trous noirs. En 2017, Rainer Weiss, Barry C. Barish and Kip S. Thorne ont reçu un prix Nobel pour leur contribution décisive, au sein de la collaboration LIGO, à la détection d'ondes gravitationnelles. Cette détection a notamment confirmé l'existence des trous noirs de masse stellaire. En 2020, Roger Penrose a été distingué pour ses travaux théoriques sur les trous noirs en tant que prédiction de la théorie de la relativité générale, tandis que Reinhard Genzel et Andrea Ghez ont été récompensés pour avoir apporté des preuves observationnelles de la présence d'un trou noir supermassif au centre de notre galaxie. Ces prix illustrent la reconnaissance par la communauté scientifique de l'importance de la physique des trous noirs.

Par ailleurs, la collaboration Event Horizon Telescope a publié en 2019 la première observation directe, dans le domaine millimétrique, de « l'ombre » d'un trou noir. Cette « photographie » a suscité un intérêt considérable dans le monde. La communication autour de cet évènement a été particulièrement soignée, l'image ayant été dévoilée lors de six conférences de presse simultanées en Amérique, en Asie et en Europe. Elle a ensuite fait la une de presque tous les quotidiens du monde. Au-delà de l'intérêt scientifique de cette découverte, sur lequel nous reviendrons plus loin, cette couverture médiatique souligne l'intérêt également porté par le grand public aux trous noirs. Ceux-ci semblent occuper une place à part dans l'imaginaire collectif et la culture populaire, tant leurs propriétés défient l'intuition.

Considérer les trous noirs comme des objets astrophysiques n'est pas allé de soi. Si la théorie de la relativité générale a passé avec succès une grande quantité de tests expérimentaux, on a longtemps pensé que les trous noirs n'en étaient pas une prédiction physique. À partir des premiers résultats d'Albert Einstein et Karl Schwarzschild en 1915, il a ainsi fallu plusieurs décennies pour qu'un changement de paradigme s'opère, et que les trous noirs soient pris au sérieux malgré leurs propriétés extrêmes. Comme l'a dit Edwin Salpeter alors que l'identification de la source Cyg X-1 était encore controversée, dans bien des cas, l'existence d'un trou noir doit désormais être considérée comme « l'hypothèse la plus conservatrice ». C'est justement de ce changement de paradigme qu'est né le domaine de l'astrophysique relativiste.

Afin de déterminer la nature d'une source lumineuse dans le ciel, il est nécessaire de comprendre comment cet objet interagit avec son environnement et de modéliser précisément la radiation qu'il émet. C'est l'un des thèmes principaux de l'astrophysique relativiste que de comprendre « à quoi ressemble » un trou noir, en fonction de son environnement. Loin d'être des « aspirateurs cosmiques », avalant tout ce qui se trouve autour d'eux sans laisser aucune trace observable, les trous noirs sont à l'origine des sources parmi les plus lumineuses de l'Univers observable. En particulier, ils sont capables de produire des jets de plasma pouvant atteindre des vitesses relativistes et s'étendant sur des milliers d'années-lumière. Les progrès à venir en instrumentation astronomique promettent de fournir plus de détails sur leur environnement, et donc de mieux contraindre les modèles théoriques existants. Cette thèse s'insère dans cet effort d'interprétation des observations en provenance de trous noirs, au moyen de simulations numériques. Les motivations observationnelles et théoriques de cette thèse sont décrites dans la partie I.

Conventions and notations

In this thesis, I work in Gaussian-cgs units. Vectors appear in boldface, and the nabla vector differential operator is noted ∇. I use the Einstein summation convention for implicit summation over the values taken by an index appearing in an upper position and a lower one. For brevity, in in-line math mode, I write a fraction as a/bc instead of $a/(bc)$. In special relativity, the timelike coordinate is noted $x^0 = ct$ and has the dimension of a length.

Red blocks contain technical developments on side concepts or derivations of results from the main body of the text. They can be omitted at first reading.

I provide below a list of the notations and values (in cgs units) for the physical constants which will be used throughout the text.

c	Speed of light in vacuum	$2.998 \times 10^{10}\,\mathrm{cm}\cdot\mathrm{s}^{-1}$
m_e	Electron mass	$9.109 \times 10^{-28}\,\mathrm{g}$
m_p	Proton mass	$1.673 \times 10^{-24}\,\mathrm{g}$
e	Electron charge	$4.803 \times 10^{-10}\,\mathrm{g}^{1/2}\cdot\mathrm{cm}^{3/2}\cdot\mathrm{s}^{-1}$
h	Planck constant	$6.626 \times 10^{-27}\,\mathrm{erg}\cdot\mathrm{s}^{-1}$
G	Gravitational constant	$6.674 \times 10^{-8}\,\mathrm{cm}^3\cdot\mathrm{g}^{-1}\cdot\mathrm{s}^{-2}$
$M_\odot$	Solar mass	$1.988 \times 10^{33}\,\mathrm{g}$

Part I

Introduction: context and motivations

- Mais c'est quoi tout ça ?! C'est pas un rapport ! Y a combien de pages ?
- Non mais ce qu'il y a de bon, c'est qu'il y a de tout là-dedans. Comment fabriquer un avion... L'itinéraire des bus de nuit... En y regardant bien, il doit même y avoir la recette de la tarte tatin.

- What kind of report is this?! How many pages are there?
- On the bright side, you can find a bit of everything in there. How to build an airplane... The routes of night buses... Upon closer look, there must even be a tarte tatin recipe.

LANZAC & BLAIN, Quai d'Orsay, chroniques diplomatiques, tome 2

Chapter 1

A brief history of black holes

The concept of black holes is one of a kind. The discovery of most concepts in physics is driven by experiments, while theory lags behind, working to design a unified and consistent view of the observed phenomena. On the opposite, black holes were conceived in purely theoretical ways. This solution behaved in such peculiar ways that it was almost immediately rejected and deemed unphysical. It took several decades, some very bright minds and a fantastic wealth of observations to comprehend the astrophysical relevance of black holes. In this section, my goal is to describe this realization from an astrophysical point of view. The material in this section relies on the historical accounts in the books by Thorne (1995) and Begelman & Rees (2009), which I can only recommend.

1.1 A theoretical concept

1.1.1 The premises

A prehistoric version of black holes had been suggested in the 1780's by John Michell and Pierre-Simon de Laplace. In view of the corpuscular theory of light that held sway at the time, they figured that according to Newton's laws, an object could become so massive that the escape velocity from this object would be greater that the speed of light (Michell, 1784). Back then, nothing prevented one from increasing indefinitely the compactness of an object. These proto-black holes were dubbed "dark stars", but the concept did not survive long after Thomas Young's wave description of light was published and generally accepted. Thereafter, the concept remained dormant until the beginning of the $20^{\rm th}$ century, when Albert Einstein published his theory of general relativity.

The first hint of black holes appeared in the first solution of Einstein's equations for a spherically symmetric object, derived by Karl Schwarzschild in 1915 (Schwarzschild, 1916). This concrete solution revived the idea that the propagation of light could be modified in a gravitational field. At the time, scientists were inclined to study the consequences of this solution in the weak-field limit, that is, provided the masssive object had a radius much larger than the Schwarzschild radius $r_{\rm s} = 2GM/c^2$ (with M the mass of the object). However, they did not take seriously its extreme predictions in the case of a very compact object with a radius close to $r_{\rm s}$. There was a sort of tension between the acceptance of the general theory of relativity, due to its first experimental successes, and the widespread reluctance to take what was then called "Schwarzschild singularities" seriously. Even Albert Einstein wrote that "Schwarzschild singularities do not exist in physical reality" as the conclusion of a paper (Einstein, 1939).

In the late 1920's, the dominant astrophysical puzzle was white dwarfs, which have masses close to the solar mass $M_\odot$ but much smaller radii. In particular, this means that white dwarfs are made of extremely dense matter, much denser than what could be found on Earth. Because they are so dense, there must be a source of pressure much stronger than thermal gas pressure counteracting the strong squeeze of gravity, but the origin of this pressure was unknown. Subrahmanyan Chandrasekhar showed, by inserting special relativistic

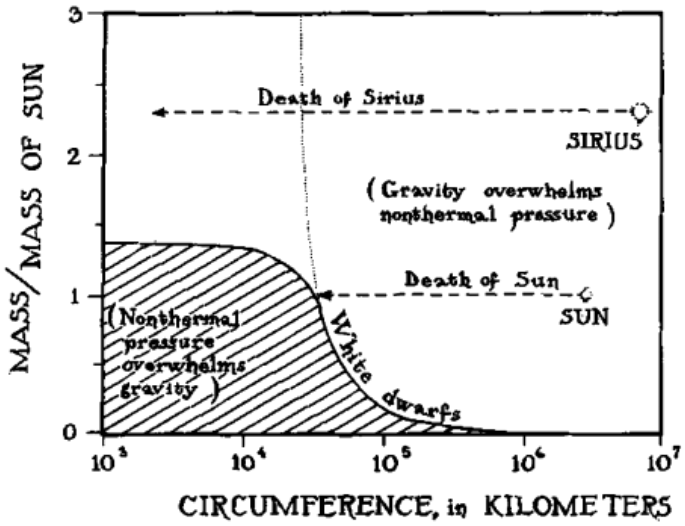


FIGURE 1.1: Stellar mass-radius diagram after Chandrasekhar's discovery. As they age, stars move leftward in the course of their lifetimes. Stars with mass greater than $M_{\rm C}$ cannot reach equilibrium as a white dwarf. Figure taken from Thorne (1995).

arguments to the newly founded quantum mechanics, that the counteracting force was due to the degeneracy pressure of electrons. More importantly, he showed that if the star got too massive, even this strong degeneracy pressure could not withstand the pull of gravity, thereby placing an upper limit $M_{\rm C}$ on the mass of white dwarfs (Chandrasekhar, 1931). Using realistic equations of state for white dwarf matter, he found the upper limit (Chandrasekhar, 1935)

$$M_{\rm C} \simeq 1.4\, M_\odot. \tag{1.1}$$

As a regular star ages, it radiates and cools down. As a consequence, in order to maintain the pressure/gravity balance, stars must shrink gradually. When a star less massive than $M_{\rm C}$ completes burning its nuclear fuel, so that the thermal pressure cannot resist gravity at all, it shrinks suddenly until the degeneracy pressure takes on (see Fig. 1.1). But what happens to stars initially more massive than $M_{\rm C}$ was the question that no one really wanted to ask. The mental block at the time was the impossibility to conceive that stars could implode.

1.1.2 The fate of massive stars

Neutron stars have been hypothesized by Franck Zwicky (and independently by Lev Landau) in the 1930's, shortly after the discovery of the neutron, to account for the enormous power emitted in supernovæ (Baade & Zwicky, 1934). They provided a convenient way out of the fact that stars could shrink indefinitely at the end of their lives. If objects could be mostly made up of neutrons, then a combination of neutron degeneracy pressure and nuclear force could take over the electron degeneracy pressure at greater masses, and prevent the collapse of the star to a black hole. So the question amounted to determining whether neutron stars existed at all, and if they did, whether they could be arbitrarily massive. This problem was investigated by Richard Tolman, George Volkoff and Robert Oppenheimer, who derived the famous TOV equation for the radial structure of a star in general relativity. They were able to show that there is indeed a maximum mass $M_{\rm TOV}$ for neutron stars, although the exact value, of the order of several solar masses, depends on unknown nuclear forces and remained uncertain (Oppenheimer & Volkoff, 1939). On top of that, Robert

Oppenheimer and Hartland Snyder performed idealized modeling of the collapse of spherically symmetric, presssureless stars, in full general relativity (Oppenheimer & Snyder, 1939). They studied the implosion both from the point of view of an external observer and one falling with the star, which allowed scientists to apprehend the collapse of a dying star. This study was the first to show the physical consequences of the Schwarzschild singularity at $r_{\rm s}$ (for example, the extreme time dilation that makes the star appear frozen to an observer at infinity); it pointed that physics do not break down as the observer passes the Schwarzschild radius.

This led to the conclusion that nothing could prevent stars initially more massive than $M_{\rm TOV}$ from collapsing to black holes, unless there should exist yet another "graveyard" between black holes and neutron stars. The latter idea was dismissed by John Wheeler, who worked out the equation of state of cold matter for a very wide range of densities, bridging the gap between neutron stars and white dwarfs (Wheeler, 1966). They combined it with the TOV equation to compute the relation between the mass and radius of cold stars (see Fig. 1.2). Although the equation of state of neutron stars is probably much more complicated than that of cold matter, there did not seem to be a way to produce arbitrarily massive neutron stars. Nothing can prevent the implosion of massive stars.

> Even today, the neutron star equation of state is not fully understood. It is still possible to show that the mass of a neutron star cannot be above $\simeq 3M_\odot$ (Rhoades & Ruffini, 1974). More details about neutron star models can be found in the textbook by Shapiro & Teukolsky (1983).

The last loophole to prevent the formation of black holes would be if the star somehow ejected a significant fraction of its mass ahead of the complete collapse. After World War 2, computing capabilities and understanding of the physics at play in stars had both made such progress that realistic simulations of collapsing stars could be carried out (Colgate & White, 1966). These confirmed Oppenheimer's idealized calculations (Podurets, 1964; May & White, 1966), making the idea of black holes even firmer on astrophysical grounds. Besides, on the theoretical field, analytical extensions of the Schwarzschild metric were found, highlighting the fact that the Schwarzschild singularity was really just a coordinate singularity, so that both the interior and the exterior of the object can be described jointly (Finkelstein, 1958; Kruskal, 1960). It was shortly thereafter that John Wheeler coined the term "black holes", which quickly took hold in the community.

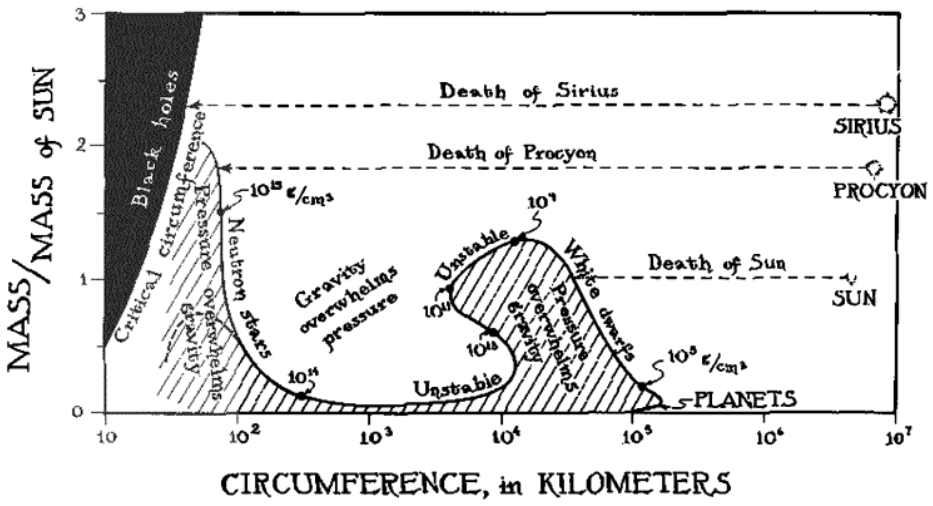


FIGURE 1.2: Extended stellar mass-radius diagram, after it was realized neutron stars have a maximum mass. Figure taken from Thorne (1995).

> About the fate of massive stars, Eddington said during a talk to the Royal Astronomical Society in 1935 that "the star will go on radiating and radiating and contracting and contracting, until, I suppose, it gets down to a few kilometers radius, when gravity becomes strong enough to hold in the radiation, and the star can at last find peace". All in all, these were quite prescient words, although Eddington remained skeptical about black holes until the end.

1.1.3 Academical black holes

In the 1960's, everyone was finally convinced that black holes were perfectly plausible objects which did not break the laws of physics. After struggling with the complexities of general relativity, theoreticians realized that black holes were conceptually quite simple objects. Let us present a short list of the main results of this decade.

- Another solution, describing a rotating black hole, was found by Kerr (1963) (see Chap. 4).
- Black holes inevitably enclose a singularity, no matter the degree of symmetry of the initial collapse (Penrose, 1965).
- Energy can be extracted from a rotating black hole (see Sec. 4.2.2).
- Black holes can pulsate, and by doing so they radiate away any inhomogeneity in the spacetime metric (Price, 1972a,b). Such pulsations extract gravitational energy from the black hole, but they are stable and die out with time (Press & Teukolsky, 1973).
- The Schwarzschild and Kerr metric are unique solutions (see Sec. 4.1.2).
- The no-hair theorems were proved (see Sec. 6.1.2): black holes radiate away any degree of freedom other than their mass, electric charge and angular momentum.

It is worth noting that the acceptance of black holes was a "collective" effort, in that various areas of physics were involved. The development of this concept is intertwined with the flourishing and understanding of quantum mechanics, nuclear forces, gas dynamics... These theories were developed at roughly the same time and contributed to making the idea of black holes plausible.

1.2 The astrophysical age: stellar-mass black holes

So far, black holes had remained abstract, idealized entities, in the hands of theoreticians. This was all well and good, but the question of their observability remained. By nature, black holes cannot be observed as they do not cast any light (putting Hawking's radiation aside). The only hope was to detect them through their impact on their close environment. The first scientists to lay out astrophysicallly realistic scenarios where black holes could leave an observable signature were Edwin Salpeter and Yakov Zel'dovitch (Salpeter, 1964; Zel'dovich, 1964) in the late 1960's. Among other ideas, they figured that binary systems composed of a star and a black hole could shine in X-ray. They were the first to point out that accreting systems could present very high luminosities.

1.2.1 Accretion

Accretion means the accummulation of gas toward a central massive object by the pull of gravity (Longair, 2011). It is one of the most efficient ways to convert mass into energy. Let us consider the accretion of matter onto an object of mass M and radius R. A proton

initially at rest at infinity has a kinetic energy $E = GMm_{\rm p}/R$ when it reaches the surface of the object. If the accretion rate is $\dot{m}$, then the rate of energy deposited onto the surface of the object is

$$L = \frac{GM\dot{m}}{R} = \frac{r_{\rm g}}{R}\dot{m}c^2, \tag{1.2}$$

where $r_{\rm g} = GM/c^2$ is the gravitational radius of the object. Traditionally, we write

$$L = \eta\dot{m}c^2, \tag{1.3}$$

with η the radiative efficiency of conversion from kinetic energy to radiation. At first sight, L is all the larger as the compacity of the object increases, that is, as R approaches $r_{\rm g}$. General relativity states that the most compact objects are black holes, with R lying between $r_{\rm g}$ and $2\,r_{\rm g}$ depending on their spin. Neutron stars come quite close to this limit, with $r_{\rm g}/R \simeq 0.15$.

Now, if the object is a white dwarf or a neutron star, then the kinetic energy of the accreted material is deposited to the stellar surface, heating it. In that case, Eq. (12.5) gives the total luminosity of the object. The radiative efficiency of these objects is almost precisely equal to their compactness: $\eta \simeq r_{\rm g}/R$. In contrast, black holes lack a material surface. As a result, the kinetic energy of the inflowing gas must be emitted before it reaches the event horizon. The radiative efficiency of black holes is therefore smaller than $r_{\rm g}/R$. For example, spherically symmetric, radial accretion is unlikely to produce significant emission.

Because the compact object is so small with respect to the large size of the accreting material reservoir, infalling matter almost necessarily retains a certain amount of specific angular momentum. As the gas rotates around the black hole, friction tends to flatten it in the direction of its total angular momentum, until it forms an accretion disk. Its angular momentum prevents it from falling in the perpendicular directions. In such a disk, the gas' angular velocity depends on the distance to the central object: the disk is in differential rotation. It follows that a massive particle is unable to flow toward the black hole unless it loses angular momentum. Transport of angular momentum is necessary in order for accretion to occur. This transfer of angular momentum can operate precisely because of differential rotation. The nature of this interaction causing this transfer is still a matter of debate.

The luminosity of the object cannot be arbitrarily large. If the accretion rate is too high, radiation pressure from the inner parts of the disk will be able to push back the infalling material. The upper limit on the luminosity $L_{\rm Edd}$ is obtained by balancing radiation pressure and gravitational force. Under the simplifying assumption that Thomson scattering brings the main contribution to the radiation pressure, we obtain the Eddington luminosity

$$L_{\rm Edd} = \frac{4\pi GMm_{\rm p}c}{\sigma_{\rm T}} = \frac{4\pi r_{\rm g}m_{\rm p}c^3}{\sigma_{\rm T}} = 1.3\times10^{38}\left(\frac{M}{M_\odot}\right)\mathrm{erg\cdot s^{-1}}. \tag{1.4}$$

This is an estimate of the upper limit on the luminosity of accreting systems, and it provides an order of magnitude of their typical luminosity. Note that this calculation is highly idealized: spherical accretion is assumed, and the opacity of the gas to photons is only due to Thomson scattering. Some systems can be super-Eddington, with $L > L_{\rm Edd}$, but this is more the exception than the rule.

As said earlier, the kinetic energy of the gas must be converted into radiation before it reaches the event horizon. Let us assume that the conversion into heat is very efficient ($\eta \simeq 1$), and that the disk radiates like a blackbody (in that case, it is said to be optically thick: photons can only escape from the disk at the surface). From the Stefan-Boltzmann law, we can estimate the temperature T of the inner parts of the disk:

$$4\pi r_{\rm g}^2 \sigma_{\rm B} T^4 \simeq L_{\rm Edd} \Rightarrow T \simeq 10^7 \left(\frac{M_\odot}{M}\right)^{1/4} {\rm K}, \tag{1.5}$$

where $\sigma_{\rm B} = 5.670 \times 10^{-5}\,{\rm erg\cdot cm^{-2}\cdot s^{-1}\cdot K^{-4}}$ is the Stefan-Boltzmann constant. This means that the disk emits most of its energy in the X-rays.

Much more accurate calculations can be carried out in order to obtain the structure of the disk and the dependence of the total spectrum with the parameters $\dot{m}$ or M. The accretion disk model was first studied by Lynden-Bell (1969), and then in much more detail by Shakura & Sunyaev (1973) (in flat spacetime) and by Novikov & Thorne (1973) (in the Kerr metric).

Radiative efficiency

Massive particles in circular orbits around a black hole cannot remain stable arbitrarily close to the black hole. Stable timelike circular trajectories are bound by the innermost stable circular orbit (ISCO) at $r = r_{\rm ISCO}$. For a non-spinning black hole, the ISCO is located at $r_{\rm ISCO} = 6\,r_{\rm g}$. Gas in an accretion disk settles into circular orbits around the compact object, and slowly gains energy as it drifts inwards. The maximum binding energy $mc^2 - E$, where E is the energy of the particle, is reached when the particle is at $r_{\rm ISCO}$. This is unlike a material compact object, where all the gravitational potential energy of the particle would be retrieved at the surface. For a Schwarzschild black hole, the maximum radiative efficiency is (Shapiro & Teukolsky, 1983)

$$\eta_{\rm max} = 1 - \frac{E_{\rm min}}{mc^2} = 1 - \frac{2\sqrt{2}}{3} \simeq 5.7\%. \tag{1.6}$$

If the black hole is spinning, the prograde ISCO can get much closer to the event horizon, thus increasing the maximum radiative efficiency. In that case, the maximum efficiency can reach $\simeq 40\%$ (Shapiro & Teukolsky, 1983).

1.2.2 X-ray astronomy

Until the 1950's, the astronomers' view of the sky was basically limited to the optical wavelengths, although radio astronomy was starting to flourish (more details in Sec. 1.3.1). The 1960's finally saw the rise of X-ray astronomy, which opened brand new perspectives and paved the way for modern astronomy (Begelman & Rees, 2009). Since X-rays cannot penetrate the Earth's atmosphere, X-ray detectors must be launched in orbit. The first X-ray detectors were launched in 1962 on a rocket by a group led by Riccardo Giacconi. They had very poor angular resolution and gathered little data, but they identified an unexpectedly bright X-ray source in the constellation of Scorpius (Sco X-1), along with a diffuse background (Giacconi et al., 1962). This came as a surprise, as everyone expected the brightest source in the sky to be the Sun. This prompted research into more resolved X-ray detectors, which were mounted on satellites this time. The first X-ray satellites were *Uhuru* in 1970 and *Einstein* in 1978. We now know that the sky is filled with unresolved, pointlike X-ray sources (see Fig. 1.3).

Different kinds of sources were spotted. Some pulsated quite regularly, while others like Cyg X-1 flickered very quickly, on timescales of tens of milliseconds (Cui et al., 1997; Pottschmidt & Konig, 1997). This is shown in Fig. 1.4(a). In the first case, the compact object is a neutron star, with a pulse period consistent with the rotation period of the star. In the second case, there is no clear time signature. However, causality constraints place

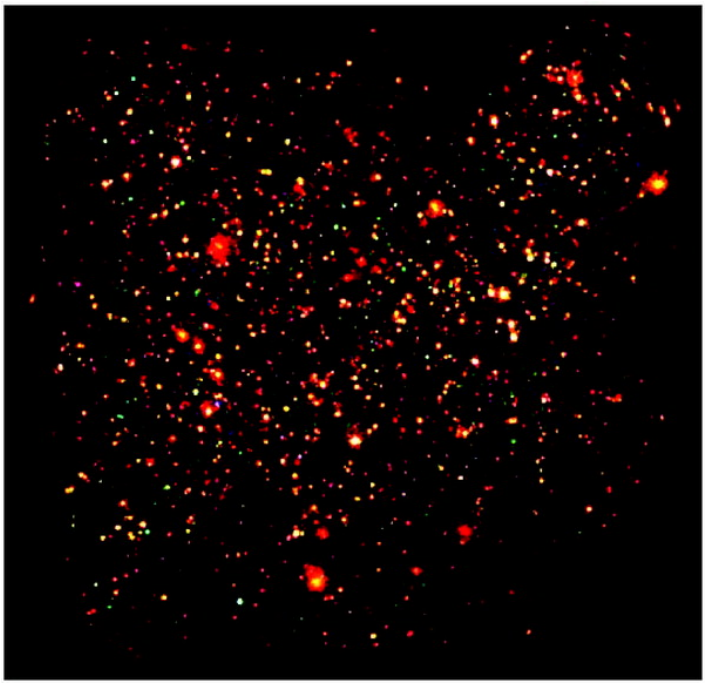


FIGURE 1.3: Image of the sky in X-ray, taken by the *XMM-Newton* space telescope. The different colors correspond to different energy intervals. Image taken from Hasinger et al. (2007).

an upper bound on size of the emitting region. For an emitting region of size R, setting relativistic boosting aside (see Sec. 2.1), significant variations in the light curve cannot occur on timescales shorter than the light-crossing time $t_{\rm g} = R/c$. In the case of a black hole of radius $R = r_{\rm g}$, this translates into

$$t_{\rm g} \simeq 10^{-5} \left(\frac{M}{M_\odot}\right) {\rm s} \simeq 10^{-5} \left(\frac{M}{M_\odot}\right) {\rm s}. \tag{1.7}$$

This is the typical timescale which can be associated with a black hole of mass M.

Cyg X-1 was the first candidate to be a stellar mass black hole, the natural product of the collapse of a massive star. Such sources, comprising a star and a compact object surrounded by an accretion disk, are called X-ray binaries. The gas in the accretion disk originates from the companion star. This source displayed variability consistent with an accreting system powered by a stellar-mass black hole, according to Eq. (1.7). However, it is *a priori* not possible to reject a neutron star as the compact object powering this source, although no pulsations are seen, on the sole basis of its flickering. In both cases, most of the energy is expected to be converted before reaching the surface, so the radiative features should be similar. How to determine unequivocally the nature of the central object from its spectral and timing properties is still an open problem (Barret et al., 1996; Titarchuk & Shaposhnikov, 2005).

In order to discriminate between neutron stars and black holes, in the absence of regular pulses, the least equivocal way is to determine independently the mass of the compact object. In most cases, this is very challenging, due to the unknown inclination of the binary or the faintness of the companion star. Yet, in the case of Cyg X-1, this was done by measuring the orbital period and speed of the companion star (Webster & Murdin, 1972; Bolton, 1972; Sowers et al., 1998), through the Doppler shift of its optical emission lines as it revolves around the compact object (see Fig. 1.4(b)). Astronomers found that the mass of the compact object was at least $3.5\,M_\odot$, which is more than the maximum mass of neutron

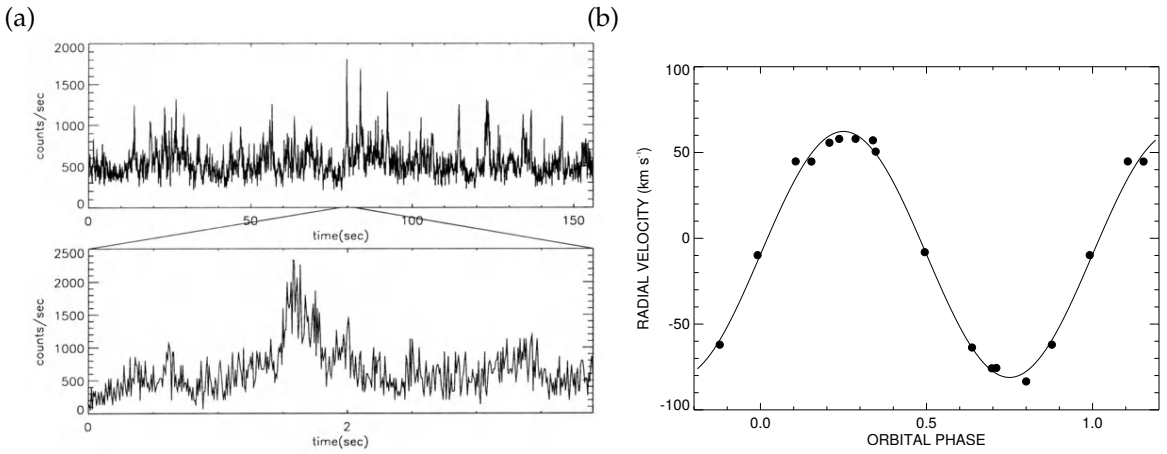


FIGURE 1.4: (a) X-ray light curve of Cyg X-1, taken by the EXOSAT X-ray satellite. Figure taken from (Pottschmidt & Konig, 1997). (b) Radial velocity of the companion star of Cyg X-1 against orbital phase. The solid line shows the best fit for a circular motion. Figure taken from Sowers et al. (1998).

stars. Hence, it became very likely that Cyg X-1 harbours a black hole. The modern estimate for the mass of the black hole has recently been improved; it was previously estimated at $\simeq 15\,M_\odot$ (Orosz et al., 2011), but a recalibration of the distance to the system changed it to $\simeq 21\,M_\odot$ (Miller-Jones et al., 2021). About 20 other X-ray binaries in our galaxy were thought to host a black hole on such basis. Remarkably, all their luminosities lie approximately below $10^{39}\,\mathrm{erg\cdot s^{-1}}$ (Corral-Santana et al., 2016), which is the Eddington luminosity with $M \simeq 100\,M_\odot$ (see Eq. (1.4)). This dynamical measurement is often hard to perform, as the secondary star can be too faint for us to detect its radial velocity.

> In the 1970's, the status of the compact object in Cyg X-1 was so uncertain that Stephen Hawking bet against Kip Thorne that the compact object was *not* a black hole.

Since then, the existence of stellar-mass black holes has also been vividly demonstrated by the detection of gravitational waves (Abbott et al., 2016). See Abbott et al. (2017) for a pedagogical interpretation of the implications of this detection. Interestingly, efforts have been made to infer the presence of an event horizon in stellar mass black holes being part of X-ray binaries. Comparing the luminosity of black-hole X-ray binaries with neutron-star X-ray binaries for similar accretion rates seems to show the following trend: black-hole X-ray binaries are less luminous than their neutron-star counterparts (Narayan & Heyl, 2002; McClintock et al., 2003, 2004). This would be explained naturally by the loss of material inside the event horizon, compared to the accreting material hitting and heating the surface of the neutron star.

1.3 The astrophysical age: supermassive black holes

A detailed history of the discovery of active galactic nuclei is given by Shields (1999).

1.3.1 Radio astronomy

We saw in the previous section how how X-ray astronomy changed our view of the sky, and made the discovery of stellar-mass black holes possible. Similarly, the discovery of supermassive black holes was due in large part to the advent of radio astronomy. The first

detection of galactic radio waves was made by Karl Janksy in 1931, in the context of transatlantic telecommunications. He noticed that the brightest sources were originating from the Milky way, rather than the Solar system (Jansky, 1933). His work was followed more than ten years later by Grote Reber, a radio amateur who looked for cosmic radio signals and published the first radio maps of the sky (Reber, 1944), shown in Fig. 1.5 (see Kellermann (1999) for a historical discussion on this pioneering work).

One of the brightest radio sources, Cyg A, was located in the constellation of Cygnus. Unfortunately, because of the large wavelengths employed, radio telescopes had very poor angular resolution. This prevented any identification of radio sources with their optical counterparts. The discovery of radio interferometry allowed significant progress in the 1950's, with the key realization that arrays of radio dishes could be used to increase the angular resolution, taking advantage of the Earth's rotation to reconstruct images. The angular resolution was improved by a factor 10 to 100 in a mere decade, and the localization of the radio source Cyg A could be probed with sufficient accuracy for the astronomers to conduct an optical survey of the region. Walter Baade was eventually able to identify Cyg A with a distant faint galaxy, with redshift 0.05 (Baade & Minkowski, 1954). The contrast between the faintness of the optical emission and the intensity of the radio emission was particularly striking. Following on this, and after another improvement of the angular resolution of their radio array, Jennison & Das Gupta (1953) showed that the radio emission from Cyg A originated from two huge lobes, symetrically disposed with respect to a central radio source (see Fig. 1.6). Comparison with the optical data showed that the galaxy coincided with the central radio source. This strongly suggested that the central galaxy was at the source of the radio emission, and pointed toward a new kind of galactic activity: galaxies do not simply emit through the combined light of their stars. These objects became known as "radio galaxies".

It had already been recognized by Ginzburg (1982) that galactic radio emission was likely due to synchrotron emission from relativistic electrons (see Sec. 8.3), due to the power-law form of the spectrum and the high degree of polarization. Inferring the distance from Earth to Cyg A, by measuring its redshift and radio luminosity, allowed to provide an estimate of the energy budget. Geoffrey Burbridge showed that the total energy content of the radio lobes was of the order of $10^{59}\,\mathrm{erg}$, with a power $\simeq 10^{44}\,\mathrm{erg\cdot s^{-1}}$ (Burbidge, 1956, 1959). Somehow, galaxies were able to deliver enormous outputs of energy into the intergalactic medium, in the form of relativistic electrons and magnetic fields. The origin of this energy was not understood at the time.

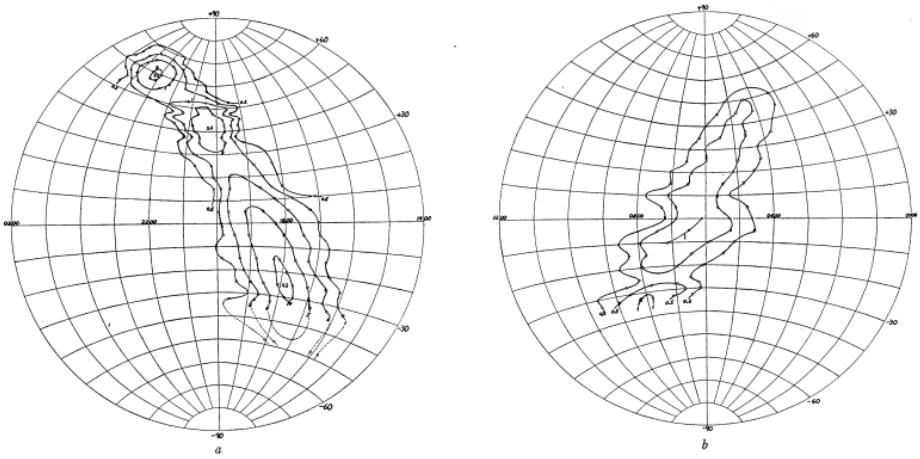


FIGURE 1.5: Lines of constant intensity in the radio sky. Figure taken from Reber (1944).

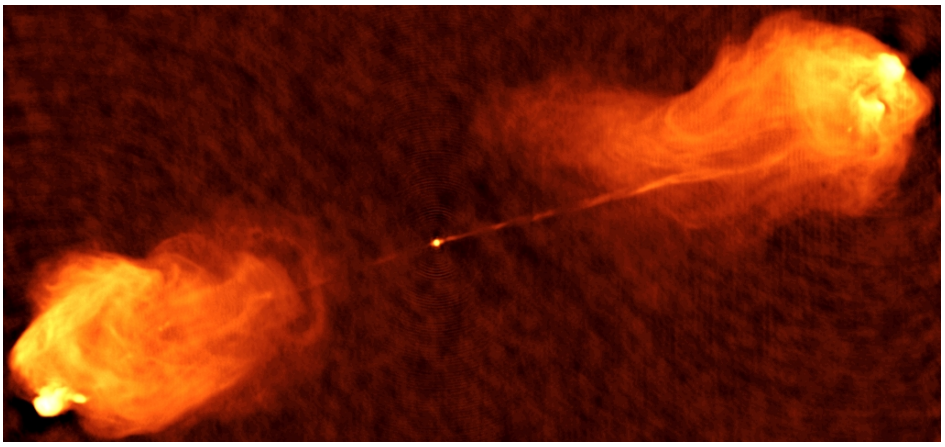


FIGURE 1.6: Radio image at 5 GHz of the source Cyg A (Carilli & Barthel, 1996). Courtesy of NRAO/AUI/NSF.

1.3.2 Active galaxies and the discovery of quasars

Optical astronomers had also made progress on their own in the studies of galaxies. They realized that galactic nuclei were unexpectedly blue and luminous. This was seen in both spiral galaxies, in the 1940's (Seyfert, 1943), but also in elliptical galaxies in the 1960's (Markarian, 1967). The spectra of these bright nuclei was not so starlike. They showed intense and broad emission lines, the Doppler broadening corresponding to much higher velocities than ordinary gas motion in a galaxy. On the opposite, galaxy spectra usually display stellar absorption lines, along with narrow emission lines from star-forming regions. Besides, the nuclear optical continuum emission was detected to be variable on timescales of days (Fitch et al., 1967). Theses galaxies were coined "Seyfert galaxies", a subclass of "active galaxies". In visible light, they looked very much like normal galaxies, but it was later realized that the total bolometric luminosity of their cores was of comparable intensity to the combined luminosity of the stars compounding it. These nuclei also showed signs of non-thermal activity in radio or X-ray. There again, this new type of activity remained mysterious.

The major breakthrough came from a joint effort of radio and optical astronomers and the discovery of quasars. Unlike Cyg A, several radio sources (3C 48 for instance) were found to be pointlike, and remained unresolved even with the increasing angular resolution of radio arrays. They were associated with optical sources characterized by strange spectra, with emission lines at unknown wavelengths (Matthews & Sandage, 1963). Then, the radio source 3C 273 was pinpointed with excellent accuracy at the location of an unusually bright optical source (Hazard et al., 1963). Maarten Schmidt recognized that although the absolute wavelengths corresponded to no known element on Earth, the relative wavelengths matched hydrogen emission lines, but shifted with a redshift of 0.16 (Schmidt, 1963). The cosmological origin of this redshift (due to the expansion of the universe, according to Hubble's law) was not immediately accepted, but no other realistic explanation came up. To detect an optical source 2 billion light-years away from Earth was unprecedented: the source 3C 273 was the brightest source ever found (see Fig. 1.7(a)). The previously found pointlike sources were also found to match this pattern, with emission lines compatible with large redshifts. Even larger redshifts were found shortly afterwards.

Those objects were named "quasars", after their first appearance as quasi-stellar radio sources. They resembled stars superficially, and had appeared on optical sky surveys much sooner, but had only been investigated closely when identified with bright radio sources. Later on, radio-quiet counterparts were discovered, and it was realized that they were much more common than radio-loud quasars (Sandage, 1965). Systematic, large-scale searches for

(a) (b)

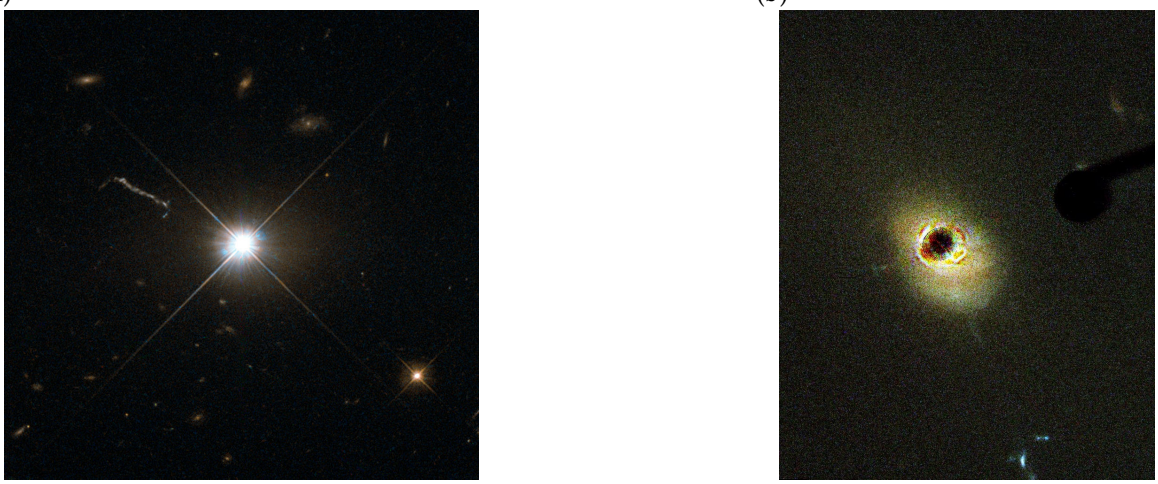


FIGURE 1.7: (a) Image from the quasar 3C 273 taken with the Hubble telescope. The diffraction spikes show that the quasar is truly pointlike. Credits: ESA/Hubble. (b) Image of the same object from Hubble Space Telescope's Advanced Camera for Surveys (ACS) (Martel et al., 2003). The ACS' coronograph blocks the light of the central nucleus, allowing to visualize the quasar's host galaxy. Credits: NASA/ESA.

quasars were performed, and the quasars' luminosity function was computed as a function of redshift z, and was found to match smoothly that of Seyfert galaxies at $z = 0$. Kristian (1973) showed that quasars events could be found in the nuclei of galaxies by imaging the host surrounding galaxy on optical plates (see Fig. 1.7(b) for modern data in the case of 3C 273), whereas Gunn (1971) proved that clusters of galaxies containing quasars had the same redshift than these quasars. These gradual discoveries allowed the astrophysical community to remove their mental block and accept the cosmological origin of quasar redshifts. They also supported the fact that Seyfert galaxies and quasars (and, in general, all active galaxies) could be unified under a common paradigm: a galaxy comprising an active nuclei. The nature of the active nuclei is the subject of the next section.

1.3.3 Identification with supermassive black holes

Aside from their huge energy outputs, active galactic nuclei (hereafter, AGN) and quasars in particular have other striking features. Their light curves in optical or X-ray display very rapid variability, up to timescales of hours. As said earlier, this places stringent constraints on the size of the emitting region (see Eq. (1.7)). AGN must be extremely compact. Besides, they are found to emit radiation over a very broad range of frequencies, from the radio band to the very high-energy gamma rays. This indicates that highly non-thermal processes are at play. Finally, AGN involve gas motions at velocities close to c, as can be deduced from the Doppler broadening of emission lines.

AGN must be powered by a central, compact object. Once again, the birth of this idea owes much to radio astronomy. Indeed, it was realized that, in powerful and resolved radio sources such as Cyg A, the central galaxy was also shining in radio. Even better, improved radio arrays showed that a thin straight line of radio emission linked the lobes to the galaxy, highly suggesting that it was the central radio galaxy which furnished energy to the radio lobes (see Fig. 1.6). Energy had to originate from the compact nucleus, and be transferred to the lobes by the means of jets carrying enormous power. From there, it became obvious that the power supply of AGN could only be of gravitational origin, by the means of accretion onto the central compact object. The requirements in luminosity and Doppler broadening

could be combined to determine the masses of AGN. This analysis was performed by (Wandel & Mushotzky, 1986) for instance. They estimated the mass from Doppler broadening of the lines, which yielded the typical velocity v of the moving emitting gas. The distance of the gas to the central object r could be obtained from the ionization parameter, so that the mass of the compact object was given by $v^2 r/G$. The inferred masses ranged from $10^6 M_\odot$ to $10^9 M_\odot$. The vast majority of AGN remains below the Eddington luminosity, with typical efficiencies of 1% to 10%. What is more, mass estimates obtained from the causality constraints of time variability (Eq. (1.7)) pointed toward masses in the same interval. No other alternative to black holes as the compact objects powering AGN could really pass the test of time: either such an object could not remain stable over millions of years (supermassive star, very dense cluster of stars), or its size would be too large to provide sufficient radiative efficiency η (see Eq. (1.3)). Lynden-Bell (1969) had put forward, as early as in 1969, that AGN could be powered by supermassive black holes, but this was largely ignored at the time. The discovery of X-ray binaries (see Sec. 1.2), and the slowly mounting body of observational evidence and theoretical arguments, along with the lack of viable alternatives, drove the community to accept the supermassive black hole model (Rees, 1984). The differences between galaxies as powerful as quasars and less luminous active galaxies is likely due to environmental factors.

Furthermore, the huge size and impressive straightness of the jets implied that the original power source had to remained steady for very long periods of time. This can be explained within the context of supermassive black holes. In the case where gas is accreting toward a spinning black hole, the angular momenta of the accretion flow and the black hole can be misaligned. Actually, the spinning black hole has the ability to warp the accretion disk, so that in the inner regions of the flow, the tilted disk lies in the equatorial plane of the black hole. This is called the Bardeen-Petterson effect (Bardeen & Petterson, 1975), a special case of frame dragging by Kerr black holes (see Chap. 4). Spinning black holes behave like stable gyroscopes. Another qualitative feature of AGN which is nicely accounted for by black holes is their excess luminosity in the optical and ultraviolet bands (Malkan & Sargent, 1982), as shown in Fig: 1.8. We have given in Eq. (1.5) a crude estimate of the temperature of the inner regions of an accretion disk surrounding a black hole with mass M. For a supermassive black hole with $M = 10^8 M_\odot$, this yields $T \simeq 10^5$ K, so that the associated blackbody

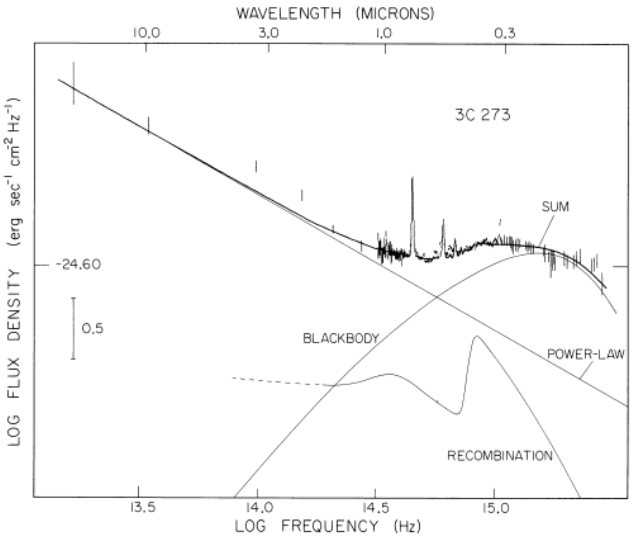


FIGURE 1.8: Optical spectrum of the quasar 3C 273, showing a fit of the continuum emission with a blackbody component in the UV. Figure taken from Malkan & Sargent (1982).

spectrum peaks in the UV.

All in all, the black-hole paradigm is consistent with most observational properties of AGN. There is an extended zoology of active galaxies, and one major challenge is to unify their various features under the black-hole paradigm, by adjusting environmental factors and our line of sight toward the galaxy. However, as we will see in Chap. 2, this paradigm alone is by no means sufficient to understand the whole multi-wavelength emission from AGN. A detailed model of AGN is still lacking, and this has been a major topic of research in the last decades. Besides, unlike stellar-mass black holes, the origin and formation of supermassive black holes is still an open problem. They might have formed by accretion of matter during a billion years or so, during their active quasar phase, or by successive mergers with other black holes (Soltan, 1982; Kulier et al., 2015).

1.4 Evidence for supermassive black holes in galactic nuclei

Direct evidence for supermassive black holes was still missing. Just like stellar mass black holes, this was mainly by lack of a better alternative. Nevertheless, it became important to find more direct observations of supermassive black holes, at least in the case of nearby galaxies, to test the black hole paradigm. Since all galaxies are not active, it is interesting to know whether even inactive galaxies harbour supermassive black holes.

1.4.1 Stellar dynamics

A general idea is to probe the dynamics of stars in galactic nucleus (Begelman & Rees, 2009). Unlike stars in galactic disks, these move in random directions, just like stars in globular clusters. Their formidable kinetic energy balances the pull of gravity. If a galaxy hosts a black hole, then the stellar average speed must increase near the center. The presence of a black hole should also increase the central concentration of stars, trapped in the black hole's gravitational sphere of influence. There should be an increase in starlight near the galactic center, although this fact alone does not imply the presence of a black hole (Kormendy & Richstone, 1995). Much information can be obtained from the spectrum of the combined stellar light: the redshift of the emission lines is used to detect regular circular motion, whereas the width of the lines is used to detect random motions and measure the dispersion in velocity σ_0. The cusp in the brightness generally matches the dynamical center. From these quantities, using the collisionless Boltzmann equation for a self-gravitating system, the mass inside a spherical region of radius r can be obtained (Longair, 2011). A crude estimate of the mass M is provided by the virial theorem: $M \simeq r{\sigma_0}^2/G$. A review of this technique, and some associated caveats and preexisting assumptions, is given by Kormendy & Richstone (1995). In the end, a nearly conclusive proof that the galaxy hosts a supermassive black hole is if the mass-to-luminosity ratio increases toward the center to values which are unreachable by stellar clusters, and confined to radii too small for clusters of brown dwarfs or dead stars to be stable, with central masses of a few billion solar masses.

These methods have been pioneered by Young et al. (1978) and Sargent et al. (1978) in the case of M87. Firmer detections have followed, in M31 for example (Kormendy & Bender, 1999). This technique allowed to build an extensive list of several dozen galaxies with central massive objects believed to be supermassive black holes, bringing evidence that supermassive black holes are actually very common objects in galactic nuclei. Masses ranging from millions to billions of solar masses have been inferred. This allowed to highlight a correlation between the black hole's mass and that of the galactic bulge, but also between the black hole's mass and the velocity dispersion of the bulge component of the galaxy (Gültekin et al., 2009). This illustrates the strong influence of the black hole on its surroundings,

although the difference in scale and mass between the central object and the bulge is enormous.

1.4.2 Gas dynamics

M87 is one of the nearest elliptical galaxies, and one of the most massive in the nearby universe, located in the constellation Virgo. Stellar dynamics measurements happened to be very difficult in this specific case. A complementary approach was to detect gas flows orbiting the central object, and measure their circular velocities. Ford et al. (1994) discovered a disk of ionized gas close to the nucleus of M87 with the Hubble telescope. The jet was seen to be parallel to the disk axis, which suggested that this disk was in the galactic plane, in Keplerian rotation around the nucleus. Measurements of the gas velocity showed that the nucleus hosted an object with mass $\simeq 3 \times 10^9 M_\odot$.

A caveat of this approach is that it is often hard to find tracers close enough to the dynamical center. Besides, gas responds more strongly to non-gravitational forces (such as radiation pressure). Better angular resolution was obtained in the case of NGC 4258 using water-maser emission at $1.3\,\mathrm{cm}$, observed with the VLBA (Miyoshi et al., 1995; Moran et al., 1999). Another benefit from this approach was that water-maser lines are very narrow, so the orbital velocities of the water clouds could be determined very accurately. In this case, so many emitting clouds have been detected that the velocity of the clouds could be plotted against the distance to the center, and a Keplerian disk was found to fit the data with exquisite precision (see Fig. 1.9). The central mass was computed to be $\simeq 3.6 \times 10^7 M_\odot$. The excellent quality of the fit indicates that this mass must lie within the innermost detected cloud. The inferred density and compactness can only be consistent with a black hole. Again, the radio jets emitted by this active galaxy are found to be perpendicular to the disk plane. Unfortunately, detecting maser emission requires very specific conditions of observation (in NGC 4258, the disk is seen almost edge-on). Overall, this method is not as prolific as stellar dynamics.

Another method relies on the observations of broad iron lines in X-ray (Tanaka et al., 1995). These lines, when emitted by the inner parts of the accretion disk, look asymmetric and skewed, as a result of Doppler boosting and gravitational redshift (Fabian et al., 2000). From these, it is possible to estimate the velocity of the emitting gas and the associated emission region, giving contraints on the mass contained inside that region. Iron line diagnostics also turned out to be a powerful diagnostic to estimate the spin of black holes for instance (Reynolds, 2020). This method depends on the presence of a geometrically thin accretion disk (see Sec. 2.2.1).

Finally, the reverberation mapping technique allows to estimate the masses of AGN in which "broad line regions" (BLR) are directly observable (Blandford & McKee, 1982; Peterson & Bentz, 2006). BLR are interpreted as gas clouds lying close to the black hole, which are photoionized by its continuum light. Their velocities are quite large ($\simeq 10^6\,\mathrm{m\cdot s^{-1}}$), resulting in broad emission lines. Reverberation mapping consists in measuring the light travel time between the central object and the BLR. Because the intensity of the continuum varies on timescales of a few days, these variations are later echoed by the BLR. The measured time delay can be converted into the distance from the black hole to the BLR. Combining this with the line-of-sight velocity of the BLR, one can obtain the mass of the central object using techniques similar to those described in Sec. 1.4.1. This method is superior to the stellar dynamics measurements, in that it probes much more compact regions (about $10^3\,r_\mathrm{g}$).

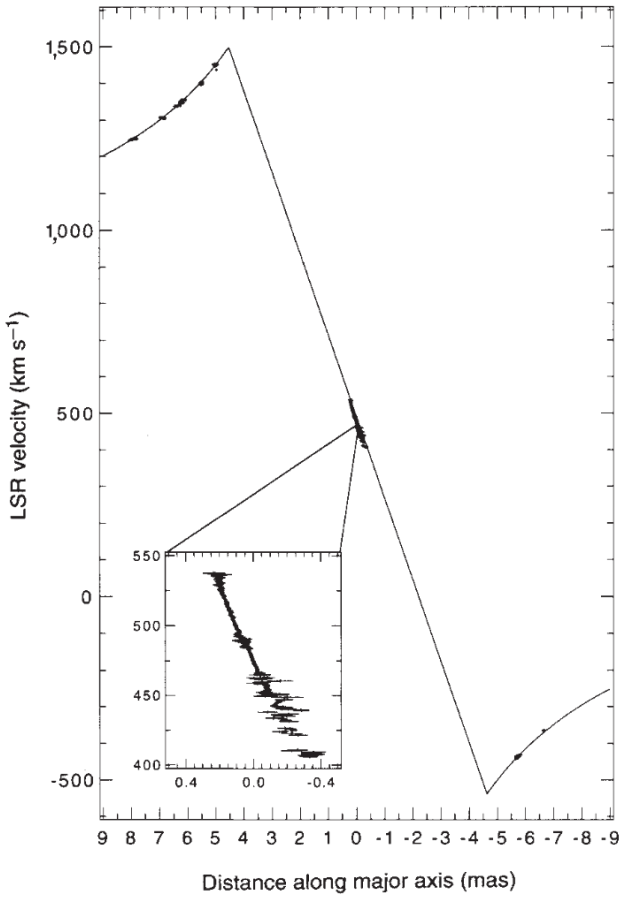


FIGURE 1.9: Line-of-sight velocity of the emitting clouds against distance along the major axis. The continuous line shows the Newtonian prediction for a central point mass of $3.6 \times 10^7 M_\odot$. Figure taken from Miyoshi et al. (1995).

1.4.3 Modern perspectives

For a long time, despite these estimates, the firmest proof of the existence of black holes remained the high mass of the companion on X-ray binaries. In the 1990's, however, observations of individual stellar trajectories in the Galactic center confirmed the presence of a supermassive black hole at the center of our Galaxy (Genzel et al., 2010). The motions and emission properties of the stars in the intense gravitational field of the black hole were confirmed by the GRAVITY collaboration (Gravity Collaboration et al., 2018b, 2020a). This will be described in more detail in Sec. 2.3.1. More recently, in 2019, the Event Horizon Telescope was able to observe the shadow a the supermassive black hole in the M87 galaxy, providing spectacular evidence in favor of supermassive black holes in galactic nuclei. More details will be given in Sec. 2.2.4.

In a relatively near future, ESA plans to launch the first space-based gravitational wave observatory: LISA (Amaro-Seoane et al., 2017). This mission will broaden the range of measurable gravitational wave frequencies. LISA should have a high sensitivity from 10^{-4} Hz to 10^{-1} Hz. This is complementary with the current ground-based observatories LIGO and Virgo, which are sensitive to higher frequencies, and are consequently more suited to the study of stellar-mass black hole mergers. In particular, LISA should be sensitive to the late stages of supermassive black holes inspirals, which will give insight into the formation and evolution of supermassive black holes (Katz et al., 2020).

Chapter 2

Multi-wavelength observations of supermassive black holes

In this section, we review some important properties of AGN cores and jets, with a focus on two of the most studied examples: M87* and Sgr A*. The main questions arising from these observations are summarized in Sec. 2.5.

Active galaxies represent about 10% of all galaxies. Their defining characteristic is their broad spectral energy distribution (SED) (Elvis et al., 1994), from radio to gamma rays, associated with strong variability. This is very much unlike non-active galaxies. An exhaustive review of their observational properties can be found in Padovani et al. (2017). Focusing first on luminous, "regular" objects, AGN emit roughly equal amounts of energy in each frequency range. All other things being equal, their radio emission can vary by more than two orders of magnitude, defining the two classes of radio-quiet and radio-loud objects. The radio emission is of synchrotron origin. The infra-red emission can often be interpreted as thermal radiation from dust which is ionized by the UV light emitted by the nucleus (Hughes et al., 1993). Their large UV flux probably comes from an optically thick accretion disk. Besides, almost all AGN are found to emit significantly in X-ray. The primary process is thought to be inverse Compton scattering of photons emitted by an accretion disk by a population of relativistic electrons, located in a "corona" (Haardt & Maraschi, 1993), which might be powered by magnetic reconnection (Beloborodov, 2017). This X-ray emission can be reprocessed on the accretion disk. The X-ray luminosity is found to be strongly correlated with the accretion-disk emission (Lusso & Risaliti, 2016), illustrating a physical relation between the X-ray emission mechanism and the accretion disk. The gamma-ray emission will be described in more detail in Sec. 2.4. Beyond these generalities, AGN come in very different flavours and have given rise to a rich phenomenology.

> By convention, if M87 denotes a particular galaxy, M87* refers to the compact object at its core.

2.1 Observations of jets

Yet another small fraction (less than 10%) of active galaxies are radio-loud, betraying the presence of a jet (Padovani et al., 2017). All galaxies are sources of radio emission, but radio galaxies and radio quasars can outmatch regular galaxies by eight orders of magnitude in this waveband. The powerful radio emission is due to the emission of powerful jets, collimated outflows of matter and electromagnetic fields at relativistic velocities. An excellent review of observational jet properties can be found in Blandford et al. (2019).

Kinetic jet powers can be obtained, for example, by estimating the age of the radio lobes and the energy stored in them (Rawlings & Saunders, 1991; Reynolds et al., 1996). Several other techniques are described in Prieto et al. (2016). Jet powers are often found to be of the same order of magnitude than the total bolometric luminosity of the AGN. For radio-loud objects which are known to display jets, the radio luminosity traces the jet kinetic power. Indeed, a correlation between the jet power or the radio luminosity and the

bolometric luminosity (a proxy for the latter being the infra-red or the X-ray luminosity) has been found (Rawlings & Saunders, 1991; Merloni et al., 2003; Fernandes et al., 2011), indicating a physical relationship between the jet power and the accretion power. This is the accretion-ejection paradigm, a connection between the disk and the jet (this relationship extends to stellar-mass black holes in X-ray binaries). The jet power still has to be greater than its associated radio luminosity, otherwise it would quickly decelerate instead of powering radio lobes. Jets can remain very stable when propagating on distances of several Mpc. They can show complex sub-structures. For example, the jet from M87 extends up to $100\,\mathrm{kpc}$ but has structure below the parsec (see Fig. 2.2(a)). Measurements of transverse gradients in Faraday rotation in jets provided supporting evidence for toroidal magnetic field in the jet (Wardle, 2018), implying that strong poloidal currents flow through the jet.

A very important feature of jets from AGN is that the gas moves at relativistic bulk velocities. This is inferred from the two following striking observations.

- Although some jets have two symmetrical components, as the one in Cyg A, some jets are one-sided. More specifically, the counter-jet is much dimmer than the direct jet. A prominent example is the jet from M87, at any observation frequency (see Fig. 2.2(a) and 2.2(b)).
- In some jets, sub-structures seem to move at superluminal velocities (see Fig. 2.2(c)). The apparent velocity is measured using the distance to the active galaxy, estimated by the redshift if the galaxy. Apparent velocities up to $\simeq 40\,c$ have been observed.

It is unlikely that one-sidedness should be a generally intrinsic feature of jets. More likely is that this is a consequence of relativistic beaming (see the box in the next page), which enhances the intensity of radiation emitted from matter moving toward the observer at relativistic speeds. A neat confirmation of this model was provided by the realization that in radio galaxies, the direct jet shows less depolarization by the ambient medium than the counter-jet (Laing, 1988; Garrington et al., 1988). This shows that the brighter jet is always the one coming toward us, disproving the one-jet hypothesis.

Superluminal expansion is also an aberration effect, occurring when the emitting region moves at relativistic velocities with a low angle with respect to our line of sight. Relativistic motion had first been invoked by Rees (1966) to explain very high brightness temperature in some AGN, and was later confirmed by the discovery of superluminal motion in the quasar 3C 273 using VLBI (Pearson et al., 1981). With time, it became possible to make a statistical survey of superluminal sources (Cohen et al., 2007), showing that superluminal motion is a common feature of AGN, and inferring relativistic outflows (with an upper bound on the bulk Lorentz factor $\simeq 50$). These considerations have given credit to the unification schemes, namely the idea that all AGN essentially contain an accreting supermassive black hole surrounded by a torus of dust. AGN observed with different viewing angles can have very different appearences.

Special relativistic effects

Consider an emitting blob at a great distance, moving with velocity $\boldsymbol{V} = \boldsymbol{\beta} c$ and Lorentz factor $\Gamma = \left(1 - \boldsymbol{\beta}^2\right)^{-1/2} \gg 1$, ejected with some angle θ with respect to our line of sight. The apparent velocity, projected on a plane perpendicular to our line of sight, is (Longair, 2011)

$$\beta_{\text{app}} = \beta \frac{\sin\theta}{1 - \beta\cos\theta} \tag{2.1}$$

The maximum apparent velocity is $\Gamma\beta$, for $\cos\theta = \beta$. Thus, for ultra-relativistic motion and a small viewing angle, one can perfectly have $\beta_{\text{app}} \gg 1$ (see Fig. 2.1(a)).

Another important special relativistic effect is the Doppler beaming. Let us define the beaming factor

$$\delta = \frac{1}{\Gamma\left(1 - \beta\cos\theta\right)}. \tag{2.2}$$

In the following, primed quantities refer to the measurements performed by the observer, and unprimed quantities to the properties in the emitter's frame. The frequency ν of the emission is Doppler-shifted according to $\nu' = \delta\nu$, and the intensity (power per unit frequency and solid angle) is transformed as $I' = \delta^3 I$. Therefore, in the simple case where the emitted spectrum is flat, the counter-jet is dimmer than the direct jet by a factor $\left(\left(1 + \beta\cos\theta\right) / \left(1 - \beta\cos\theta\right)\right)^3$. For $\cos\theta \gtrsim \beta$, which is equivalent to having $\theta \lesssim 1/\Gamma \ll 1$, one has $1 \leq \delta \leq 2\Gamma$, whereas one has $\delta \simeq 1/\Gamma$ outside of this interval. This is illustrated in Fig. 2.1(b).

Similarly, special relativistic effects can affect estimates of the size due to causality constraints. The typical timescale in the emitter's frame $\Delta t'$ is related to the observed variability timescale Δt by $\Delta t' = \delta\Delta t$. Therefore, the size R of the emitting region is constrained to be smaller than $c\Delta t' = \delta\left(c\Delta t\right)$, relaxing the constraint on the size by a factor δ. Note that for a source at cosmological distance, with redshift z, the size estimate becomes $R \leq \delta\left(c\Delta t\right) / \left(1 + z\right)$.

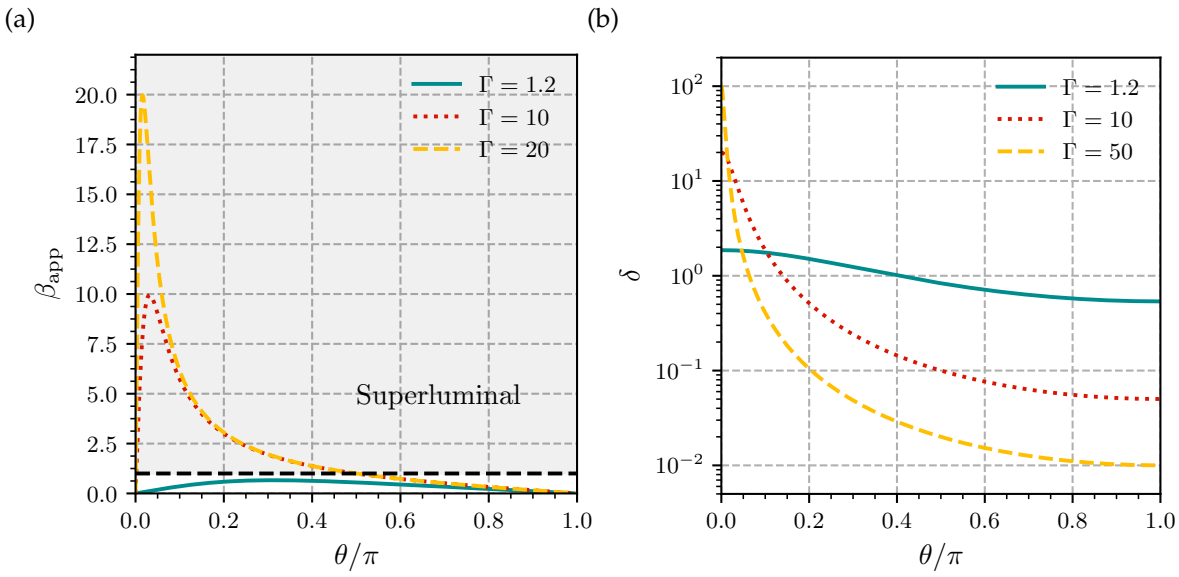


FIGURE 2.1: Apparent velocity β_{app} (a) and beaming factor δ (b) as a function of the viewing angle θ, for various values of Γ. In the left panel, the shaded region corresponds to superluminal apparent motion.

(a) (b) (c)

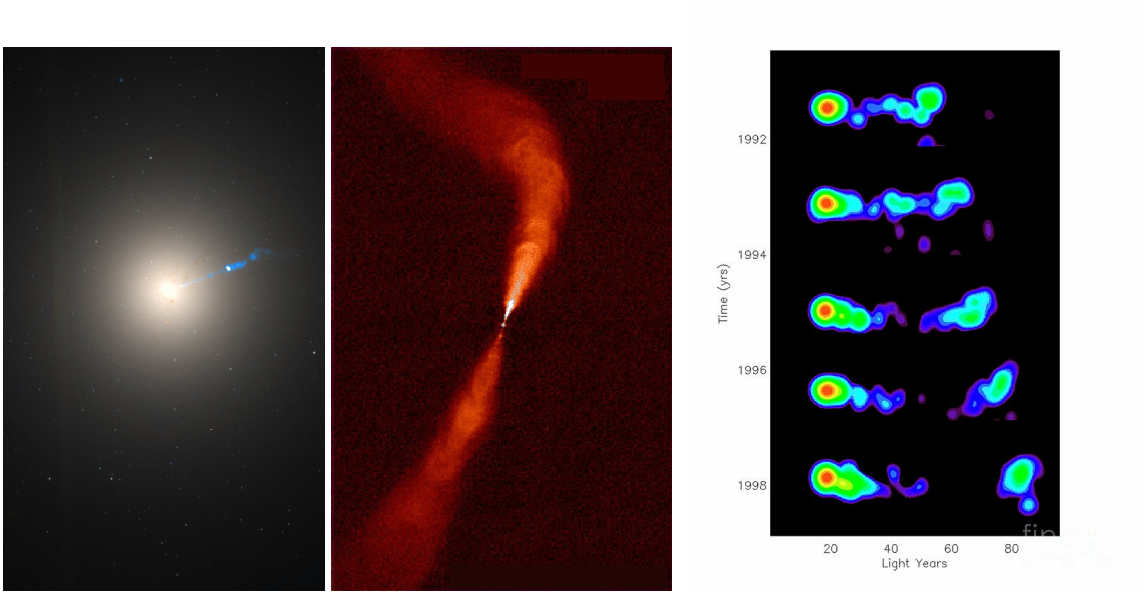


FIGURE 2.2: (a) Image of the M87 galaxy taken with the Hubble telescope. Credits: NASA/ESA. (b) Radio image of the jet emerging from the radio galaxy 3C 31. Credits: NRAO/AUI/NSF. (c) Radio image of the jet from the quasar 3C 279, over a six-year period. A blob moves at apparent superluminal speed. Courtesy of NRAO and AUI.

In the case of M87, the jet can be observed in the optical, UV, X-ray and radio-mm bands (see Fig. 2.2(a)). It is one-sided because of relativistic beaming. The jet is limb-brightened, with a spine-sheath structure: a mildly relativistic layer envelops a relativistic, accelerating core jet (Walker et al., 2018). It makes an angle $\theta \simeq 18°$ with respect to our line of sight (Mertens et al., 2016; Walker et al., 2018). The jet was tracked down to less than $7\,r_{\rm g}$ (Kim et al., 2018), suggesting that the supermassive black hole at the center must be powering the jet. Besides, proper motions of sub-structures show that the acceleration is steady from the core to about $10^5\,r_{\rm g}$, where a stationary knot is observed (HST-1) (Asada et al., 2014; Mertens et al., 2016). Downstream of this knot, the jet slowly decelerates. The large bulk Lorentz factors of jets are not reached at a short distance from the black hole, but rather result from a steady accelerating process. This suggests that the jets are initially electromagnetically dominated (Poynting-dominated), and that this electromagnetic energy is dissipated and converted into kinetic energy (Mertens et al., 2016). Such arguments require relatively low initial mass loading to explain the high terminal velocities of the jets. This argument is in line with what happens in pulsar nebulæ: an initially Poynting-dominated wind is dissipated before the termination shock (Kirk et al., 2009).

> There exists an important difference between the the two situations. In pulsar winds, the flow propagates spherically, basically unimpended, and converts most of its energy into particle thermal energy (Cerutti et al., 2020). The acceleration of the pulsar wind is due to ideal processes. In jets, the interaction of the outflow with the ambient medium cannot be neglected, and probably contributes to the jet collimation. Electromagnetic energy is mainly converted into bulk kinetic energy (Lyubarsky, 2009).

2.1.1 Launching mechanism

The launching mechanism is a question of utmost importance. Jets are also found in protostellar systems or microquasars (X-ray binaries with jets imaged in radio), advocating for a common formation mechanism. The necessary ingredients seem to be an accretion disk

threaded by an ordered, large-scale, poloidal magnetic field, and/or a central compact spinning object. Such poloidal fields have been seen to build up in the inner regions in numerical general relativistic magnetohydrodynamics (GRMHD) simulations, starting from initial poloidal field configurations (Tchekhovskoy et al., 2011) or toroidal field configurations (Liska et al., 2020). A first possibility is that the jet is powered by accretion itself, following a "magnetocentrifugal" mechanism (Blandford & Payne, 1982; Lesur, 2020). If the plasma in the accretion disk is sufficiently ionized, magnetic field lines are dragged by the differential rotation of the Keplerian disk. Considering initially vertical field lines, a radial component is induced by inwards radial accretion, which generates a toroidal component at the disk surface when this radial magnetic field is sheared. The toroidal magnetic field removes angular momentum from the disk, which is then progressively used to accelerate an outflow. As the toroidal field is exhausted in this process, a gradient of toroidal magnetic pressure forms, pushing the outflow vertically. It is the magnetic torque exerted on the disk, due to the Lorentz force, which brakes the disk and allows radial accretion by removing angular momentum. In this picture, the disk acts like a dynamo, generating a system of electric currents which extracts energy and angular momentum from the disk. More will be said about this analogy in Sec. 7.2.3. The luminosity that can be produced by this process can account for the luminosities of AGN (Livio et al., 1999). In this framework, the energy is transported in the form of a magnetically driven wind. The energy output is originally coming from the accreting material, and the luminosity of the disk cannot exceed $\eta\dot{m}c^2$.

On the other hand, independent estimates of the jet power and bolometric luminosities seem to indicate efficiencies η of order 1 (Rawlings & Saunders, 1991; Ghisellini et al., 2010; Fernandes et al., 2011; Ghisellini et al., 2014). In other words, the power of the jet exceeds $\dot{m}c^2$ (see Fig. 2.3). This can only occur if there is another source of energy. However, a non-spinning black hole has $\eta < 10\%$ and cannot accommodate such high efficiencies. Hence, there must be another energy reservoir. Besides, observations of X-ray spectra in Seyfert

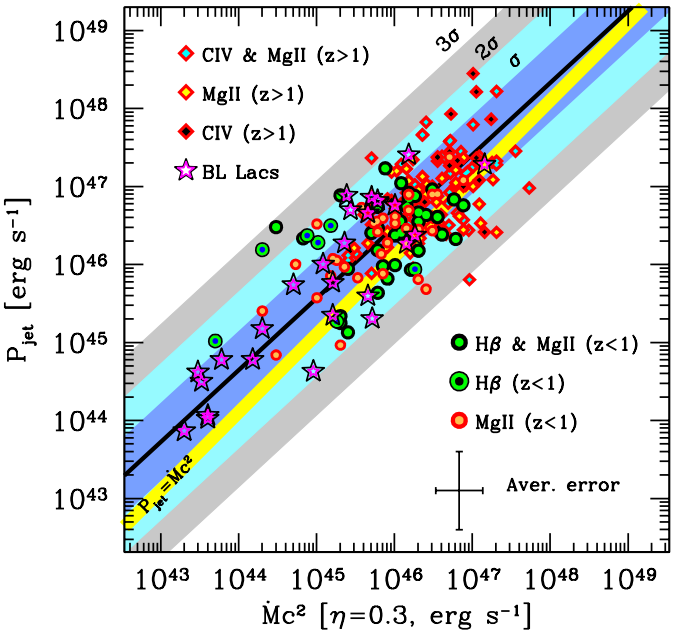


FIGURE 2.3: Jet power plotted against the accretion power for a sample of 217 blazars. The accretion power assumes an efficiency of 30%. The different symbols correspond to different ways of measuring the disk luminosity. The black line is the best fit, and the yellow stripe corresponds to equal jet and accretion powers. Figure taken from Ghisellini et al. (2014).

galaxies have also suggested that a central, compact source of X-rays should illuminate the disk (Wilms et al., 2001), which is inconsistent with a pure accretion disk model. This points toward the central object providing some extra power.

In the following, the "magnetosphere" of a compact object refers to the inner, magnetically dominated regions in the close vicinity of this object. In the case of a black hole, the magnetic field must be generated by external currents in the accretion flow and cannot be supported by the black hole itself (we will come back to this point later). It has been proposed by Blandford & Znajek (1977) that energy can be extracted from a spinning black hole threaded by a poloidal magnetic field, and embedded in a force-free magnetosphere. The electrodynamics of black-hole magnetospheres and the derivation of the Blandford-Znajek mechanism will be the main focus of Chap. 6 and 7. In this second paradigm, the jet is powered by the central compact object rather than the disk. *A priori*, this would mean that the accretion disk and the jet would have different energy sources, decoupling them in some sense, and contradicting the aforementioned correlation between the jet and accretion power. But the extraction of energy requires a magnetic field, which can be generated by dynamo, so that accretion ultimately provides the necessary means for the rotational energy to be extracted. Such a jet would initially be Poynting-dominated, which is consistent with the picture of a jet slowly converting its electromagnetic energy.

Currently, it is still a very lively debate whether jets are disk-driven or black hole-driven. Finding a correlation between the spin of the black hole and the power of the jet would help, but measurements of black-hole spins are still very uncertain. Another idea would be to probe the matter content of the jets. The Blandford-Znajek mechanism is electromagnetically dominated, with initially just enough charges to screen the electric field. The charge loading of the jet is expected to be the result of electron/positron pair creation by photon-photon annihilation. Henceforth, conversion of electromagnetic energy would also result in a mass loading of the jet with fresh pairs. Conversely, in the Blandford-Payne model, the mass of the jet should mainly be carried by protons. Unfortunately, this question is not settled.

One approach is to estimate the kinetic power of the jet. At fixed jet power and bulk Lorentz factors, a change in the matter composition can change the required plasma density in the jet by two orders of magnitude. Constraints imposed by the radio luminosity then allow to discriminate between the two. Using such arguments, Reynolds et al. (1996) made the case for an electron/positron jet in M87, whereas Celotti & Fabian (1993) rather favored an electron/proton jet. The latter would require that the population of synchrotron-emitting electrons should have a very high cut-off at low energies ($\gamma_{\rm min} \gtrsim 100$) otherwise the jet would have too much kinetic energy. On the other hand, in the case of an electron/positron jet, the pair spectrum can extend down to mildly relativistic energies ($\gamma_{\rm min} \gtrsim 1$). Unfortunately, the lowest energy leptons are not directly observable, because their radiation is self-absorbed. Another approach was taken by Wardle et al. (1998), who noted that the strong linear polarization observed in jets limits the amount of Faraday rotation suffered by the synchrotron emission, which is mainly due to low-energy electrons. This provided a way to determine the low-energy cut-off of the emittting population and discriminate between the two models. They measured the fraction of circularly polarized radiation, presumably produced by Faraday conversion of linear polarization. In the case of the quasar 3C 279, they found that the distribution of radiating particles extended to $\simeq 1$, implying an electron/ positron jet. Hirotani et al. (1999) confirmed this prediction by kinetic arguments similar to those of Reynolds et al. (1996). However, the debate is still open, and there are sound arguments against pure electron/positron jets (Sikora & Madejski, 2000). Obviously, it is very likely that both processes operate. For instance, the dense magnetically driven wind of electron/proton plasma could be responsible for the collimation of the tenuous pair plasma jet. This is the two-flow model originally advocated by Sol et al. (1989) and Henri & Pelletier (1991).

2.2 The case of M87*

2.2.1 Low-luminosity AGN

The M87 galaxy, first mentioned in Sec. 1.4.2, hosts one of the most massive known black holes, of several billion solar masses. It is a prominent radio source. What is more, it shows spectacular activity, with a one-sided jet visible in optical and radio, and non-thermal emission from the nucleus. But despite its huge mass and activity, M87* radiates at a low rate with respect to its associated Eddington luminosity. Its bolometric luminosity is $L_{\rm M87} \simeq 10^{42}\,{\rm erg\cdot s^{-1}}$ (Owen et al., 2000; Prieto et al., 2016), so that we have $L_{\rm M87}/L_{\rm Edd} \simeq 10^{-6}$ for $M \simeq 10^9 M_\odot$. This issue was first raised by Fabian & Canizares (1988). One of the reasons invoked to explain this low luminosity is an accretion rate much smaller than the Eddington accretion rate $\dot{m}_{\rm Edd}$, defined as

$$\dot{m}_{\rm Edd} = \frac{L_{\rm Edd}}{\eta c^2}. \tag{2.3}$$

For a black hole with radiative efficiency $\eta \simeq 10\%$, this represents the maximum accretion rate, beyond which the luminosity is larger than the Eddington luminosity. In the case of active galaxies, accretion should at least occur by interstellar gas falling toward the central black hole. The minimal accretion rate should thus be given by the Bondi accretion rate (see the block below).

Bondi accretion

The simplest model of accretion onto a compact object of mass M is the Bondi accretion (Bondi, 1952), which assumes spherical accretion of a uniform fluid at rest at infinity, in a hydrodynamical description. With these hypotheses, up to a numerical factor of order 1, accretion should occur at the Bondi accretion rate, given by

$$\dot{m}_{\rm B} = \frac{\rho G^2 M^2}{c_{\rm s}^3}, \tag{2.4}$$

where ρ is the mass density at infinity, and $c_{\rm s}$ the sound speed at infinity. Accretion becomes supersonic from the Bondi radius $r_{\rm B} = 2GM/c_{\rm s}^2$. The Bondi radius is generally considered as the boundary of the accretion flow: below this radius, the gravitational pull is stronger than the adiabatic expanding motion of the medium.

Chandra observations of M87 in X-ray were able to probe the gas temperature and density profiles at the Bondi radius $r_{\rm B} \simeq 230\,{\rm pc}$. The X-ray emission from this hot gas is thought to be thermal Bremsstrahlung. The Bondi accretion rate was estimated at $\dot{m} \simeq 0.1 M_\odot/{\rm year} \simeq 10^{-4}\dot{m}_{\rm Edd}$ (Di Matteo et al., 2003). This is the rate of mass that is *a priori* supplied to the accretion flow. With a canonical radiative efficiency of $\simeq 10\%$, this would yield a luminosity of $10^{44}\,{\rm erg\cdot s^{-1}}$, which is still much higher than the actual bolometric luminosity of M87.

These results prompted the community to believe that the radiative efficiency of the accretion flow was smaller than expected. They searched for accretion flow solutions departing from the standard picture of geometrically thin, optically thick disks, which can only be applied successfully to luminous sources accreting close to their Eddington accretion rates. A huge literature has been devoted to sub-Eddington radiatively inefficient accretion flows (RIAF), prime examples of which being advection-dominated accretion flows (Narayan & Yi, 1994, 1995; Reynolds et al., 1996), adiabatic inflow-outflow solutions (Blandford & Begelman, 1999) or convection-dominated accretion flows (Quataert & Gruzinov, 2000b). See Yuan & Narayan (2014) for a review on hot accretion flows. The underlying principle is that in

usual geometrically thin accretion disks (Shakura & Sunyaev, 1973), the accreting material can cool efficiently, so that all energy liberated by dissipative interactions can be radiated locally. In the opposite case, the gravitational potential energy must be stored as thermal energy, and the cooling cannot keep up with the heating (Quataert, 2003). This results in very hot accretion flows, which have a tendency to expand vertically until they are geometrically thick. Pressure forces become more important and the flow is no longer Keplerian; accretion occurs almost spherically. Thermal energy is advected toward the horizon.

It is also possible that all the mass accreting through the Bondi radius does not end up in the black hole's close environment. This means that the accretion rate must be considered as a function $\dot{m}(r)$ of radius, decreasing with decreasing r. As it happens, the Bondi estimate neglects the angular momentum of the infalling gas, which can very well be ejected by outflows beforehand. That way, the mass accretion rate can be much lower than the initial lower bound $\dot{m}_{\rm B}$. In that case, the plasma is so underdense that it is unable to convert its gravitational potential energy efficiently. So little interactions occur that the flow cannot become an optically thick disk, so that RIAF models should apply. In any case, determining the actual accretion rate onto the compact object is crucial in the prospect of discriminating between the different types of RIAFs. Numerical simulations of rotating RIAFs show that the hypothesis of an extremely low accretion rate is closest to the truth: little mass available at large radii (*i.e.* at the Bondi radius) actually accretes onto the black hole (Stone et al., 1999; Stone & Pringle, 2001).

In the case of M87, estimates of the accretion rate toward closer regions than the Bondi radius could be obtained recently. Prieto et al. (2016) find an accretion rate in quiescence $\dot{m} \simeq 10^{-4} M_\odot/\mathrm{year}$, using the spectral energy distribution (SED) of M87, whereas Kuo et al. (2014) constrain the mass accretion rate to be less than $10^{-3} M_\odot/\mathrm{year}$ at a distance of $\simeq 40\, r_{\rm g}$ from the black hole, using Faraday rotation measurements in order to probe the electron density and the magnetic field component along the line of sight (as first suggested by Quataert & Gruzinov 2000a). This confirms the results of the numerical experiments mentioned previously. All in all, the accretion flow around M87 is most likely geometrically thick, optically thin, with a mass accretion rate much smaller than the Bondi estimate. Accretion rate measurements in AGN are hard to conduct; there are only a few AGN close enough for us to resolve the Bondi radius. Aside from M87, the Bondi radius could also be resolved in Sgr A* (see Sec. 2.3.2), NGC 3115 (Wong et al., 2011) and NGC 1600 (Runge & Walker, 2021). These observations all tell the same story: the mass accretion rate at the Bondi radius is much lower than the Eddington accretion rate, and the density profile below the Bondi radius shows that the accretion rate is probably even smaller at the event horizon.

2.2.2 Spectral energy distribution

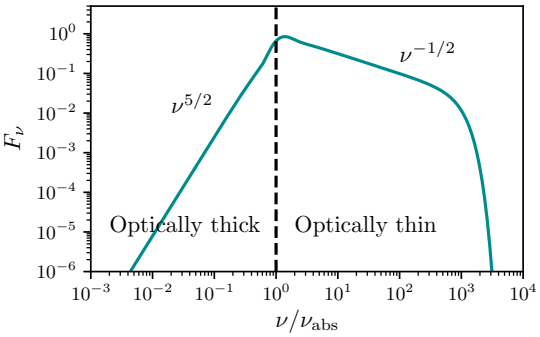
Interestingly, low-luminosity AGN lack a "blue bump" excess at UV wavelengths (Ho, 1999), which we had interpreted as the blackbody spectrum radiation from an optically thick accretion disk. This supports the fact that low-luminosity AGN are surrounded by geometrically thick accretion flows. In low-luminosity AGN, due to their very high brightness temperature, it was realized that the radio core emits synchrotron self-absorbed emission (see the block in the next page) (Blandford & Königl, 1979; Ho, 1999).

Synchrotron self-absorbed radiation

If synchrotron emission is emitted by a population of a power-law distribution of relativistic leptons, with the number density of leptons of Lorentz factor γ being $\propto \gamma^{-p}$, the spectrum of the combined synchrotron emission also follows a power-law, so that the SED reads $F_\nu \propto \nu^{-(p-1)/2}$ (Rybicki & Lightman, 1986). In the optically thin regime, this emitted spectrum equals the observed spectrum. However, this cannot be the case for abitrarily low frequencies. The synchrotron emission process necessarily comes with its inverse absorption process: synchrotron self-absorption. At low energies, as a photon propagates through the synchrotron-emitting magnetized lepton plasma, it may be absorbed by this reverse process. Therefore, all emitted synchrotron radiation does not reach the observer.

One can show that the synchrotron absorption cross section increases with increasing wavelength λ (*i.e.* decreases with increasing frequency ν). For a given size of the source, at sufficiently low energies, the synchrotron emission is actually completely "thermalized", and the source is in an "optically thick" regime. As the observing frequency increases, photons coming from deeper and deeper regions are able to escape, until the optically thin regime is reached and we recover the power law distribution $\propto \nu^{-(p-1)/2}$.

In this optically thick regime, one can show that the energy spectrum of the source follows the power-law $F_\nu \propto \nu^{5/2}$ (Longair, 2011). The spectrum rises steeply at low frequencies, until the mean free path of the synchrotron photons becomes greater than the size of the source at frequency $\nu_{\rm abs}$. The typical radio spectrum of synchrotron self-absorbed sources with a power-law electron distribution is shown in the figure on the right for $p = 2$.



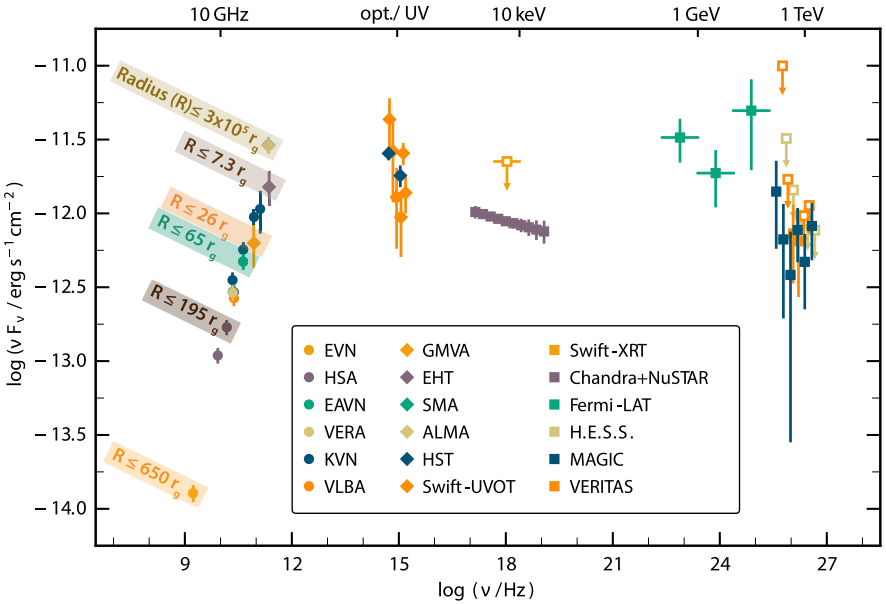


FIGURE 2.4: Simultaneous SED of the core region of M87. Note that for the millimeter and radio emission, the size of the resolved emission zone varies with observing frequency, and is labeled on the data points. Figure taken from Algaba et al. (2021).

SEDs of the cores of low-luminosity AGN are hard to study because of contamination by the host galaxy. The one from the core of M87 is shown in Fig. 2.4 (Algaba et al., 2021). The transition between the optically thick and optically thin synchrotron regime happens for wavelengths close to $1\,\mathrm{mm}$ (frequencies close to $100\,\mathrm{GHz}$).

2.2.3 The radio core of M87

Recent advances in VLBI have allowed to study the jet with superb resolution, down to the radio core. Observations have been performed with increasing frequency, starting from the GHz range, probing the synchrotron self-absorbed emission. With increasing frequency, one transitions from the optically thick to the optically thin regime, and the photosphere becomes more and more compact. At ν_{abs}, the radio core becomes entirely optically thin. Thus, increasing the frequency allows to probe regions that are closer and closer to the event horizon of the black hole, along with a better angular resolution. A montage illustrating this fact is shown in Fig. 2.5. The structure of the radio core has been studied at various wavelengths using VLBI by Mertens et al. (2016), Walker et al. (2018) and Kim et al. (2018).

Given a model for the geometry of the jet, a central prediction from the synchrotron self-absorbed nature of the emission is that the location of the radio core should depend on the observing frequency (Blandford & Königl, 1979). At a given frequency, the core emission corresponds to the photosphere of the synchrotron self-absorbed emission, so that the central engine is shifted from the radio core. This shift should decrease at increasing frequency, transitioning to an optically thin regime, and the radio core should move toward the jet's upstream side, to the central engine. This core-shift effect was observed by Hada et al. (2011), and is represented in Fig. 2.6. In particular, Hada et al. (2011) showed that the core position converges at increasing frequency, so that it should coincide with the central engine at millimeter wavelengths. See Algaba et al. (2021) for more recent observations.

2.2.4 The shadow of M87

We have come to the conclusion that the emitting region close to the event horizon of the black hole could be imaged with VLBI, provided observations were performed at millimeter wavelengths and with sufficient angular resolution, with sufficiently long baselines. Those two objectives were reached by the Event Horizon Telescope (EHT) Collaboration. The supermassive black hole M87* is one of the largest known and happens to be close to us.

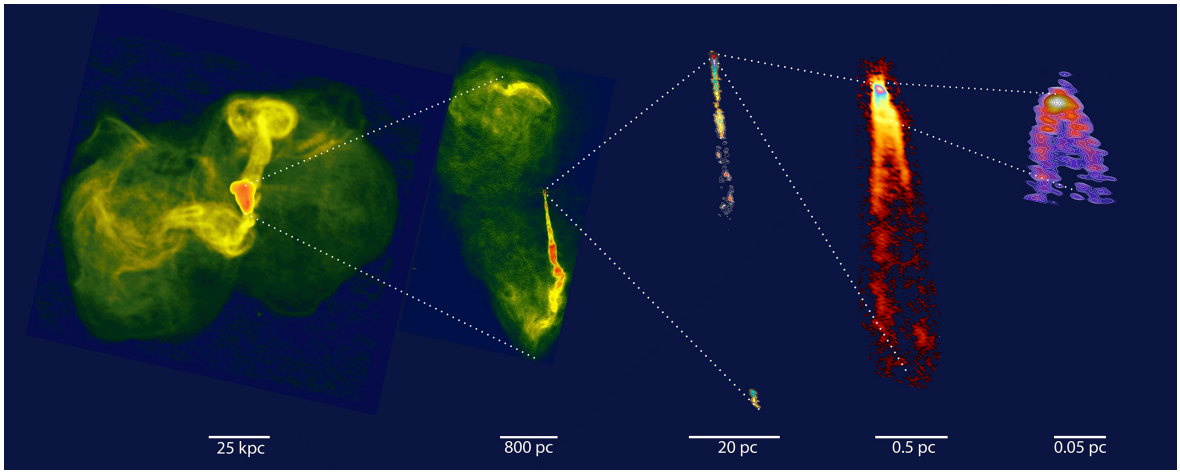


FIGURE 2.5: Images of the jet from M87 at increasing frequency, from the outer lobes to the smallest scales, taken at $\lambda = 90\,\mathrm{cm}$, $20\,\mathrm{cm}$, $20\,\mathrm{cm}$, $7\,\mathrm{mm}$ and $3\,\mathrm{mm}$ respectively. Image taken from Blandford et al. (2019).

(a) (b)

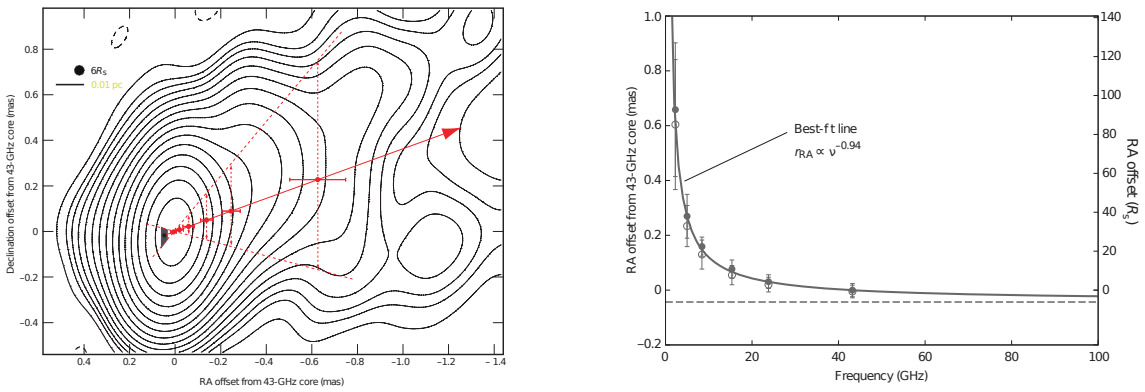


FIGURE 2.6: (a) Position of the radio core for increasing frequency, superimposed on the emission map at 43 GHz. (b) Right ascension offset from the core at 43 GHz, as a function of frequency. The dashed line shows the asymptotic behavior of the best fit. Figures taken from Hada et al. (2011).

Consequently, its angular size is the largest in the sky, which made it an ideal target for the EHT observation campaign. They gathered an array of telescopes throughout the globe observing at 1.3 mm, allowing them to reach an angular resolution of $\simeq 20\,\mu$as (Event Horizon Telescope Collaboration et al., 2019a).

At the event horizon scale, general relativistic effects become important in determining the properties of the emission. If a black hole is embedded in an optically thin accretion flow, it is expected that the black hole should cast a "shadow", such that a central dark region would appear in the image formed by an observer at infinity (Luminet, 1979; Falcke et al., 2000). Light emitted by the flow is gravitationally lensed by the black hole. The shape of this shadow is not expected to depend much on the black hole's spin or the details of the accretion flow (see Sec. 14.3.2 for more details). In the simple case of a Schwarzschild black hole with radius $r_s = 2r_g = 2GM/c^2$, surrounded by a spherically symmetric accretion flow, the shadow comes with a bright emission ring, located at

$$r_{ph} = \sqrt{27} r_g \simeq 5.2 \frac{GM}{c^2}. \tag{2.5}$$

This expression is derived in the Appendix 14.A. Modern mass measurements by stellar and gas dynamics were in a slight tension, with respective estimates of $6.6 \times 10^9 M_\odot$ (Gebhardt et al., 2011) and $3.5 \times 10^9 M_\odot$ (Walsh et al., 2013). Combined with the distance to M87 $D \simeq 16.8$ Mpc, this yields an expected angular size $\theta_{ph} = 2r_{ph}/D$ of the emission ring seen from Earth between 20 and 40 µas. Thus, provided the mass of M87* lies closer to the higher estimate than the lower, the EHT has sufficient resolution to image the black hole shadow.

The image obtained by the EHT is shown in Fig. 2.7(a). The central compact radio source of M87 is resolved, and shows a circular ring of emission with a diameter (42 ± 3) µas, which encompasses a depression in brightness. The asymmetry of the ring is likely due to Doppler boosting of material orbiting the black hole. The resolution is too low to make strong tests of general relativity or accretion flow models (see Event Horizon Telescope Collaboration et al. (2019b) for a discussion and Vincent et al. (2021) for images of hot accretion flows in alternative metrics). The inferred mass is $M = (6.5 \pm 0.7) \times 10^9 M_\odot$, closer to the stellar dynamics estimates (Event Horizon Telescope Collaboration et al., 2019c). This image makes a strong case for a supermassive black hole at the core of M87. In 2021, the EHT collaboration released polarized data (Event Horizon Telescope Collaboration et al., 2021a,b), a significant linear polarization fraction being expected due to the synchrotron nature of the

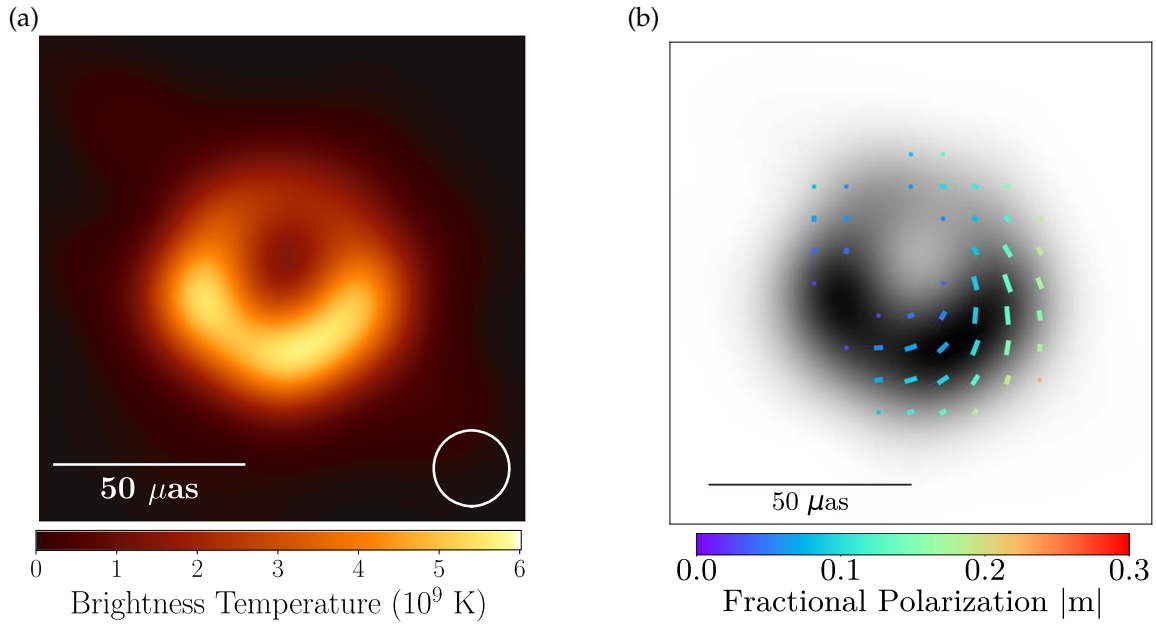


FIGURE 2.7: (a) Image of M87* from the observations of April 11, 2017, in units of brightness temperature. The circle shows the angular resolution. Image adapted from Event Horizon Telescope Collaboration et al. (2019a). (b) Same image in gray scale, with ticks showing the direction of the electric vector position angle. The tick color indicates the amplitude of the fractional linear polarization, whereas the length indicates the amplitude of the linear polarization. Image adapted from Event Horizon Telescope Collaboration et al. (2021a).

emission. This showed that the magnetic field structure in the emitting region is predominantly poloidal, implying that magnetic fields should be dynamically important. They were also able to obtain estimates of the electron density, mass accretion rate and magnetic field $B \simeq 10\,\mathrm{G}$ in this region.

2.3 The center of the Milky way: Sgr A*

2.3.1 Astrometry at the Galactic center

Obviously, the closest galactic nuclei is at the center of the Milky Way, located in the constellation Sagittarius at a distance of $8.2\,\mathrm{kpc}$ from Earth, and it has been studied extensively at all wavelengths. Since our Solar System lies in the Galactic plane, the optical and UV light from the Galactic center is obscured by gas and dust. However, we can still get a plain view in radio, infra-red and X-ray. The Galactic center turns out to be a crowded place, with a dense nuclear cluster of stars, radio-emitting filaments and plumes of X-ray, all of this embedded in the large scale Galactic bulge. At the dynamical center of the Milky Way (determined with infra-red measurements) lies a compact and variable radio source dubbed Sgr A*, which shows practically no intrinsic motion (Reid et al., 1999; Reid & Brunthaler, 2004, 2020). Gas dynamics (Lacy et al., 1980) and stellar dynamics (Genzel et al., 1996; Haller et al., 1996) suggested the presence of a supermassive black hole of about $3 \times 10^6 M_\odot$, but the bolometric luminosity of Sgr A* was extremely low, casting doubt on the black hole hypothesis.

Eventually, progress in adaptive optics made it possible to track the motions of individual stars close to the Galactic center (Ghez et al., 2000; Gillessen et al., 2009). This showed that the motion of the most central stars seemed to be caused by a high concentration of mass whose location is coincident with the radio source Sgr A* (see Fig. 2.8(a)). One star (S2), in particular, showed a highly elliptical orbit with a short orbital period ($\simeq$ 15 years)

(a) (b)

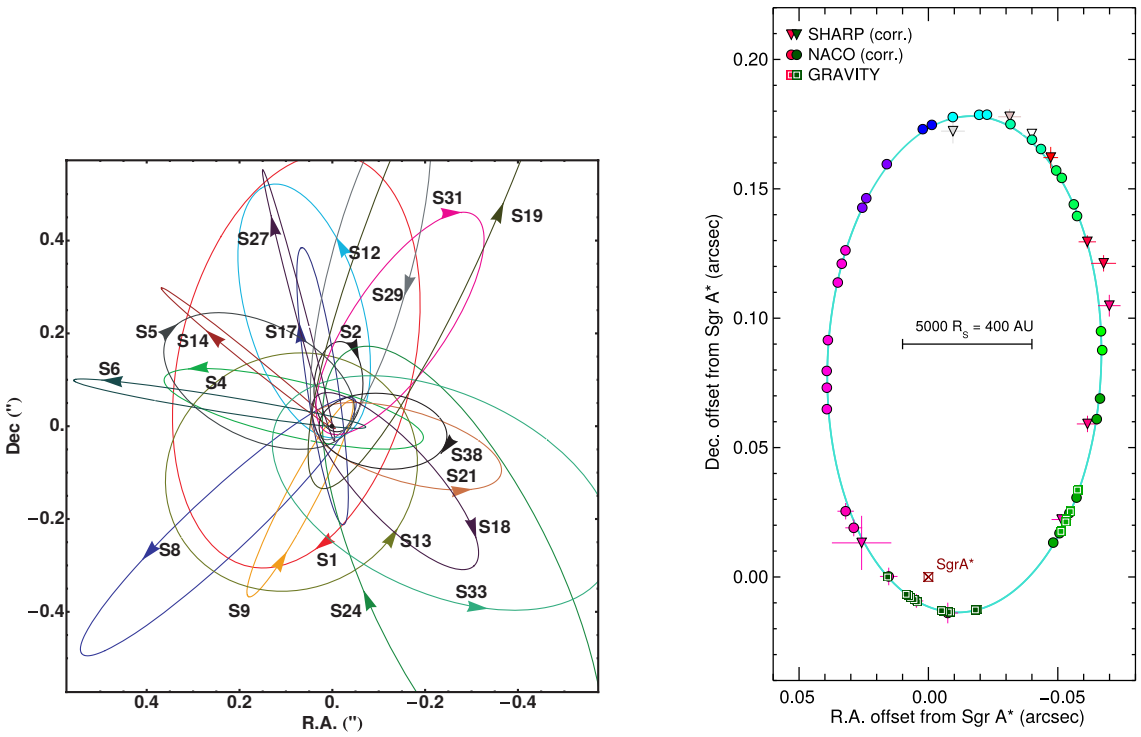


FIGURE 2.8: (a) Stellar orbits in the vicinity of Sgr A*. Image taken from Gillessen et al. (2009). (b) Projected orbit of the star S2 and position of the radio source Sgr A*. Image adapted from Gravity Collaboration et al. (2018b).

observable on a lifetime scale, as can be seen in Fig. 2.8(b) (Schödel et al., 2002). The distance of closest approach of this star to Sgr A* was measured to be $5 \times 10^{-4}\,\mathrm{pc} \simeq 200\,r_\mathrm{g}$. The low luminosity of the nucleus, its high mass and compactness are strongly indicative of the presence of a supermassive black hole. New instruments, such as GRAVITY, have used interferometry techniques in addition to adaptive optics to improve the resolution even more and probe general relativistic effects in the S2 orbit (Gravity Collaboration et al., 2018b, 2020a). An exhaustive review of the search for a supermassive black hole at the center of the Milky Way can be found in Genzel et al. (2010).

2.3.2 Quiescence

The material accreted by Sgr A* probably originates in the stellar winds of central massive stars. The bolometric luminosity of Sgr A* is $L_{\mathrm{SgrA*}} \simeq 10^{36}\,\mathrm{erg\cdot s^{-1}}$. The emission is predominant at radio and infra-red wavelengths (Genzel et al., 2010). For a black hole with mass $M \simeq 4.3 \times 10^6 M_\odot$, this means that the ratio of its luminosity to its maximum Eddington luminosity is extremely low: $L_{\mathrm{SgrA*}}/L_{\mathrm{Edd}} \simeq 10^{-8}$. This classifies Sgr A* as an (extremely) low-luminosity, almost quiescent AGN (see Sec. 2.2.1). Just like in the case of M87*, accretion through the Bondi radius was estimated using X-ray data, yielding $\dot{m} \simeq 10^{-6} M_\odot/\mathrm{year}$ (Baganoff et al., 2003). With this mass accretion rate and a canonical radiative efficiency $\eta \simeq 10\%$, the expected luminosity is much higher than the actual luminosity. Measurements of Faraday rotation by Marrone et al. (2007) and Goddi et al. (2021) restricted the accretion rate to even lower values (around $10^{-8} M_\odot/\mathrm{year}$). Again, the angular momentum of the inflowing gas and outflows prevents all the gas accreting through the Bondi radius to make it to the black hole.

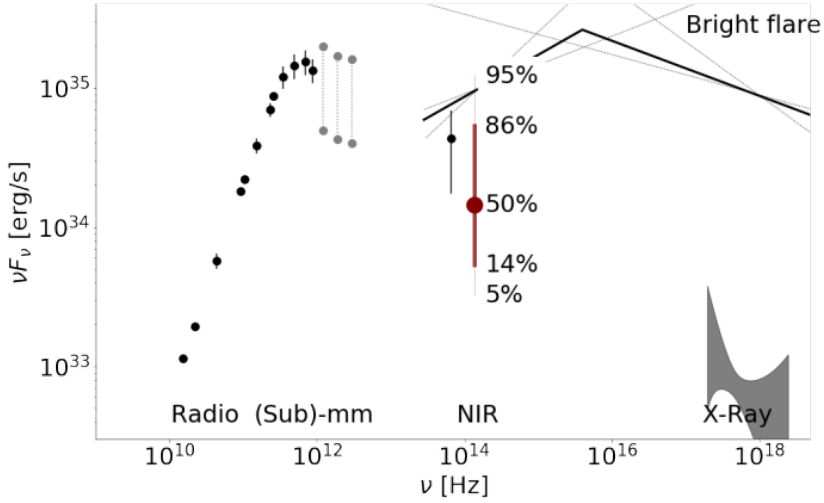


FIGURE 2.9: SED of Sgr A*. In X-ray, the quiescent and flaring emission are both shown. Image taken from (Gravity Collaboration et al., 2020b).

The SED of Sgr A* is shown in Fig. 2.9; it is typical of low-luminosity AGN and similar to that of M87*. Despite its faintness and the interstellar extinction, Sgr A* is detectable at all times in radio, infra-red and X-ray. The X-ray emission has a spatial extent of roughly the Bondi radius. The radio spectrum is due to synchrotron self-absorbed emission emitted by thermal electrons; it peaks at sub-mm wavelengths and is partly linearly polarized.

Observations of Sgr A* are made harder by interstellar scattering, which needs to be corrected. Radio interferometry at millimeter wavelength has resolved the radio core: it is extremely compact, of the scale of the event horizon (Doeleman et al., 2008). At these wavelengths, just like in the case of M87*, the regions closest to the event horizon are probed. The low luminosity, mass and compactness of Sgr A* have been used to infer the presence of an event horizon (Broderick et al., 2009). If the central compact object possessed a hard material surface, matter accreting from larger radii would deposit all its energy on the surface, which would produce unseen blackbody radiation peaking in the infra-red. Besides, a preliminary, incomplete version of the EHT array observed polarized emission from Sgr A* (Johnson et al., 2015). They measured significant polarization fractions remaining stable in time, suggesting the existence of ordered magnetic fields in the vicinity of Sgr A*, with intensity $\simeq 10\,\mathrm{G}$. Given the mass $M \simeq 4.3 \times 10^6 M_\odot$ of the supermassive black hole, its angular size in the sky is $\theta \simeq 20\,\mu\mathrm{as}$, which makes it of comparable size as M87*. Sgr A* was also a target of the 2017 EHT observation campaign; however the variability of Sgr A* is much more pronounced than that of M87*. As a result, it is harder to interpret the recorded data, taken with exposure times $\gtrsim r_\mathrm{g}/c$.

2.3.3 Variability and flares

Sgr A* shows stochastic variability in radio and infra-red, with fluctuations described by a red noise. More importantly, it exhibits strong flares, where the intensity can increase by factors of 10 to 100. X-ray outbursts were observed by Baganoff et al. (2003) and Neilsen et al. (2013), with a frequency of one flare per day on average. Infra-red flares are routinely observed (Eckart et al., 2006), at a rate of four flares per day. They are generally highly linearly polarized, and the polarization angle has been observed to swing significantly in the course of a flare (Trippe et al., 2007). Notably, almost every flare in X-ray has a counterpart in infra-red (Hornstein et al., 2007; Dodds-Eden et al., 2009). Some infra-red flares have no X-ray counterpart, due to a much better sensitivity in infra-red. Ponti et al. (2017) report

the first simultaneous detection of the spectra of the infra-red and X-ray emissions (it is denoted as the "bright flare" state in Fig. 2.9). The spectral slopes in both wavebands strongly support a synchrotron origin of the emission, with a cooling break. Typical rising times can be as short as minutes, both in infra-red (Dodds-Eden et al., 2009) and X-ray (Barrière et al., 2014). Note that the light-crossing time of Sgr A* is $r_g/c \simeq 20\,\mathrm{s}$. The radio emission is largely unaffected during flares.

The GRAVITY collaboration made a breakthrough when they detected the clockwise circular motion of a flaring hot spot around Sgr A*, shown in Fig. 2.10(a) (Gravity Collaboration et al., 2018a). For the first time, infra-red flares could be spatially and temporally resolved. Further analysis (Gravity Collaboration et al., 2020c,d) confirmed the following features of these flares.

- All flares are consistent with a circular orbit of the centroid, at a distance of approximately $\simeq 9\,r_g$ from the black hole.
- The low Doppler beaming in the course of the orbit indicates that the orbital plane is seen almost face-on, with an inclination $\simeq 140°$ (the motion is clockwise).
- The emitting region must be compact, with diameter less than $\simeq 5\,r_g$, otherwise it would have been quickly destroyed by Keplerian shear.
- The polarization angle rotates continuously along the orbit (see Fig. 2.10(b)). This indicates that the magnetic field at the flaring region has a strong poloidal component. This conclusion is similar to the subsequent EHT observations of M87* (Sec. 2.2.4): to retain a poloidal component in the toroidal sheared velocity flow, the magnetic field should be dynamically important.

The lack of sub-mm counterpart implies that flares result from synchrotron non-thermal radiation emitted by transiently accelerated electrons, to Lorentz factors $\simeq 10^3$ (Dodds-Eden et al., 2009). The physical mechanism which powers these flares is still unknown, although some scenarios have been suggested in the context of MAD simulations (Ripperda et al., 2020; Dexter et al., 2020; Porth et al., 2021). The innermost flares could also result from a

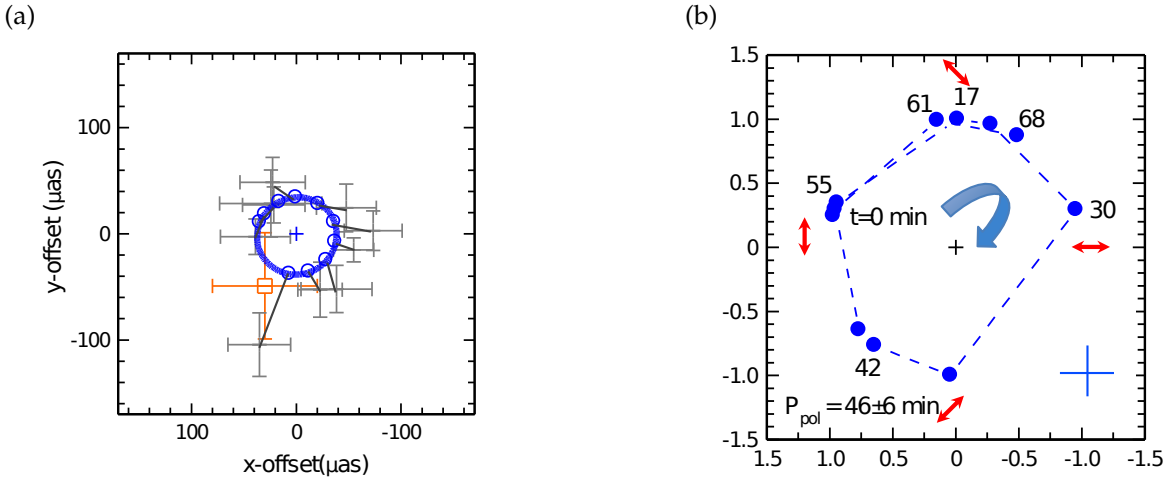


FIGURE 2.10: (a) Projected orbit of the flare centroid on the sky (grey crosses) and position of Sgr A* (orange cross). The best fit for circular motion is shown (solid blue line). Adapted from Gravity Collaboration et al. (2018a). (b) Evolution of the polarization state of the flare in the plane of normalized Stokes parameters. The red arrows show the polarization directions. Image taken from Gravity Collaboration et al. (2018a).

connection between the disk and the central compact object. The proximity of Sgr A* and the absence of a jet make it an ideal laboratory to observe processes in the close environment of black holes. It would have been difficult to resolve these flares in infra-red or X-ray in more distant AGN.

2.4 Very high-energy radiation from AGN

2.4.1 Gamma-ray astronomy and blazars

The gamma-ray band is conventionally split between the high-energy (HE) band between $30\,\mathrm{MeV}$ and $30\,\mathrm{GeV}$, and the very high-energy (VHE) band beyond $30\,\mathrm{GeV}$. A review of gamma-ray astronomy can be found in Madejski & Sikora (2016). Gamma-ray photons initiate particle showers when interacting with the atmosphere, which can be detected at the Earth surface for photons in the VHE band. Ground-based telescopes are able to track VHE photons by measuring the Cherenkov light of the ultra-relativistic leptons produced in the shower. Current operating Cherenkov telescopes are MAGIC, HESS or VERITAS. In the HE band, the showers are too faint to be detected at ground level. Therefore, HE detectors operate in space, above the atmosphere, using electron-positron pair production to capture gamma rays. The first gamma-ray satellites were launched in 1967.

Most of the gamma-ray emission originates from our Galaxy, due to pulsars or supernova remnants for example. At higher galactic latitudes, a majority of gamma-ray sources are extra-galactic. Most of these sources were found to be coincident with radio-loud AGN, many of which exhibit superluminal motion. No radio-quiet object has been detected in gamma rays. Actually, the vast majority of AGN detected in gamma ray are blazars (Madejski & Sikora, 2016; Padovani et al., 2017), which are AGN observed with a small viewing angle, for which Doppler boosting effects are particularly large. Doppler boosting of the non-thermal emission from the jet accommodates the very high luminosities and fast variability of blazars. Their gamma-ray emission is produced by Doppler-boosted inverse Compton scattering of synchrotron photons or external photons, depending on the object (Sikora et al., 1994; Madejski & Sikora, 2016).

Some extreme variability has been observed on timescales of minutes, as in the case of the blazars PKS 2155-304, observed by HESS (Aharonian et al., 2007) (see Fig. 2.11), or Mrk 501 with MAGIC (Albert et al., 2007). During rapid flares, VHE flux variations can be of one order of magnitude. The timescales of these flares are quite smaller than the light-crossing time of the event horizon, suggesting a compact emission region. On the other hand, stringent conditions on this size are set by the condition that the gamma-ray photons must escape from the flaring region: if the zone (with size R) is too compact, the opacity is too high and gamma rays are absorbed by photon-photon annihilation. Indeed, in the reference frame of the emitting region, the optical depth to pair production reads $\tau_{\gamma\gamma} \simeq n' \sigma_{\gamma\gamma} R$, with $n' \simeq L'/4\pi R^2 c\epsilon$ the density of target photons of characteristic energy ϵ (set by the luminosity L' of the region) and $\sigma_{\gamma\gamma}$ the cross-section for pair creation. Gamma rays can escape if $\tau_{\gamma\gamma} \lesssim 1$, which is all the more difficult as R decreases. Relativistic beaming again comes to the rescue, because it relaxes the constraints on R set by the variability and on the emitted luminosity L'. To account for the variability, the energetics of the emission and the possibility for gamma rays to escape, bulk Lorentz factors greater than 50 are needed (Aharonian et al., 2007; Begelman et al., 2008). Such high Lorentz factors are problematic. For example, this is in conflict with statistical studies: in the unification scheme, the fraction of blazars with respect to the population of unbeamed sources should be equal to the fraction of solid angle for which Doppler boosting leads to such enhancement of the variability. This also contradicts the moderate superluminal motion in these blazars and the measured brightness temperatures, which favour modest Lorentz factors $\Gamma \lesssim 10$ (Henri & Saugé, 2006). As

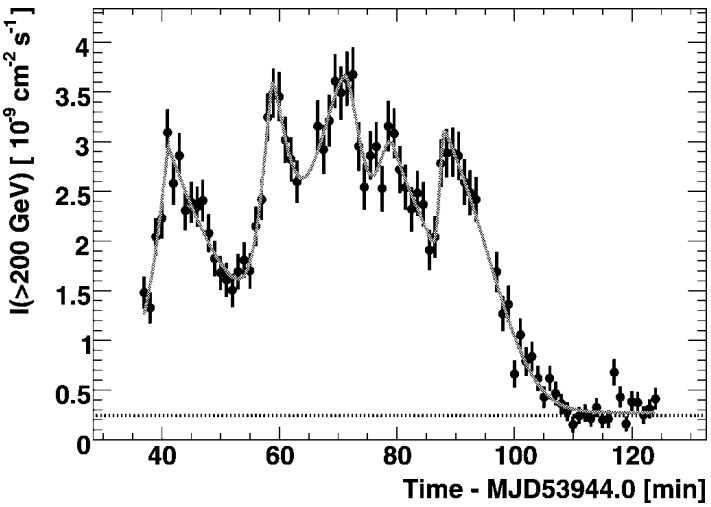


FIGURE 2.11: VHE gamma-ray light curve of the blazar PKS 2155-304 during a flare in 2006. Figure taken from Aharonian et al. (2007).

of today, this controversy is not fully solved yet, and the extreme time variability remains challenging to current theoretical models, although the estimate of the optical depth given previously is likely to be too basic.

2.4.2 Flares from radio galaxies

It came as a surprise that, among the vast number of sources detected at VHE energies, several non-blazar nearby radio galaxies were also detected: Cen A (Aharonian et al., 2009), NGC 1275 (Aleksić et al., 2012; MAGIC Collaboration et al., 2018), PKS 0625-35 (HESS Collaboration et al., 2018), IC 310 (Aleksić et al., 2014), 3C 264 (Archer et al., 2020), and (as always) M87 (Aharonian et al., 2006; Aliu et al., 2012). All of these galaxies have also been detected in the HE range by *Fermi* (Sahakyan et al., 2018), but it still remains unclear whether the HE and VHE components are distinct. A review of these observations is given in Rieger & Levinson (2018). In the unification scheme of radio-loud AGN, radio galaxies and blazars are objects of similar properties but viewed at different inclinations (Urry & Padovani, 1995). Depending on the structure of their jets, radio galaxies have been classified as FR I or FR II sources: the former are less luminous, with jets getting fainter and more distorted with increasing distance from the source, whereas the latter are straighter and more luminous, and particularly bright at their radio lobes. FR I are probably low-luminosity AGN accreting at a small rate with respect to $\dot{m}_{\rm Edd}$ (see Sec. 2.2.1). FR I galaxies are thought to accrete through a radiatively inefficient accretion flow, whereas FR II galaxies would accrete with a standard, optically thick accretion disk.

> The classification as a radio galaxy or blazar can be a bit vague, especially when the jet's inclination cannot be measured. In this case, there is some doubt as to the nature of PKS 0625-35 (HESS Collaboration et al., 2018) and IC 310 (Kadler et al., 2012), which probably lie at the frontier between the two classes (Rieger & Levinson, 2018).

Radio galaxies show misalignments with our line of sight $\gtrsim 10°$, greatly reducing the relativistic boosting effect. The dependence of δ on θ is dramatic: for $\theta = 0°$, we have $\delta = 2\Gamma$, whereas for $\theta = 10°$, the beaming factor cannot be greater than $\delta \simeq 6$. Boosted synchrotron-self-Compton radiation cannot account for this VHE emission with such low beaming factors, so that these detections might highlight new physical phenomena. Interestingly, these radio galaxies are all of a FR I type (or close to this type for PKS 0625-35

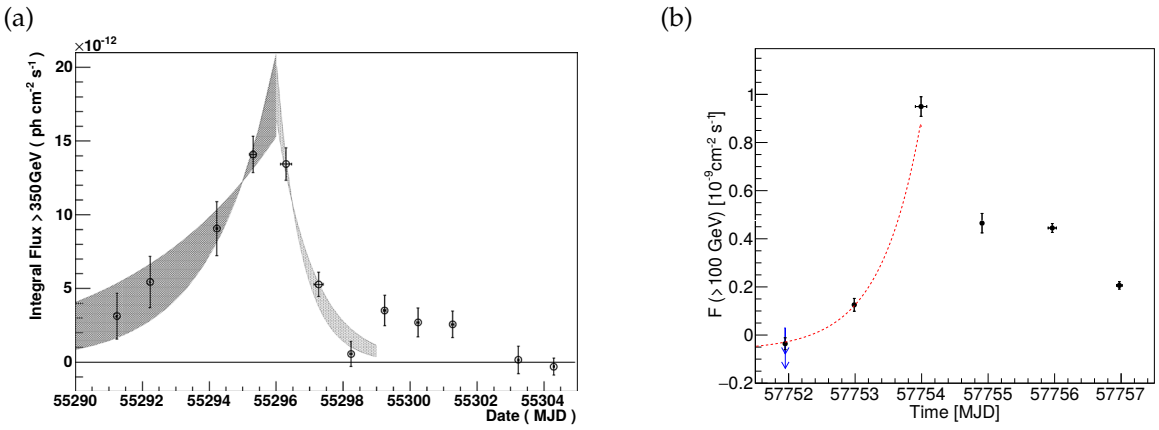


FIGURE 2.12: (a) Gamma-ray light curve of M87 during a VHE flare. The shaded regions show exponential fits for the rising and decaying phases. Figure taken from Aliu et al. (2012). (b) Gamma-ray light curve of NGC 1275 during a VHE flare. The dashed line shows an exponential fit for the rising phase. Image taken from MAGIC Collaboration et al. (2018).

and IC 310), which means that they are low-luminosity AGN. We had already M87 on the account of its low accretion rate; similarly low estimates of the accretion rate could be obtained for Cen A (Evans et al., 2004). Besides, very rapid variability has been measured in M87, NGC 1275 and IC 310 (see Fig. 2.12), although the other VHE radio galaxies are much less variable. This suggests that VHE flares are a widespread feature in radio-loud objects, not just blazars. The flares of M87 and NGC 1275 show flux rises of an order of magnitude on timescales of days, comparable to their light-crossing times r_g/c. M87 has experienced several VHE flares in the last decades. The Galactic center is also a source of GeV and TeV emission (Acero et al., 2010; Archer et al., 2014). However, the angular resolution of these observations is too poor to resolve the compact source, and these gamma rays could not be conclusively be associated with the supermassive black hole Sgr A*.

M87, in particular, has attracted considerable attention, because it can be probed at other wavelengths with excellent resolution. As a result, sub-structures in the jet can be probed in X-ray, optical and radio wavebands, while the resolution of Cherenkov telescopes is too low to allow such precise identification. It is characterized by weak beaming factor $\delta \lesssim 3$ (Dodson et al., 2006; Wang & Zhou, 2009). Its VHE spectrum extends beyond 10 TeV (Aharonian et al., 2006). Acciari et al. (2009) performed observations of M87 in all these wavelengths, and noticed a brightening of the radio core shortly after a VHE flare, whereas the flux from the knot HST-1 remained unchanged and the nucleus was in a high X-ray state (see Fig. 2.13). Such a concurrence in all wavebands was deemed unlikely to happen by chance (Acciari et al., 2009), favoring a connection between VHE flares and processes at play in the close vicinity of the central supermassive black hole M87*.

2.5 Summary

AGN produce intense activity, often in the form of powerful flares of non-thermal emission, implying that particle acceleration is triggered. The fast variability of these flares imply that this acceleration occur on very small scales. GRAVITY observations teach us that the close environment of black holes is very dynamic. At the same time, radio interferometers have been able to track the jet emitted by the supermassive black hole M87* almost down to the event horizon. The observations of VHE flares in misaligned radio galaxies have far-reaching consequences, and challenge conventional models of AGN.

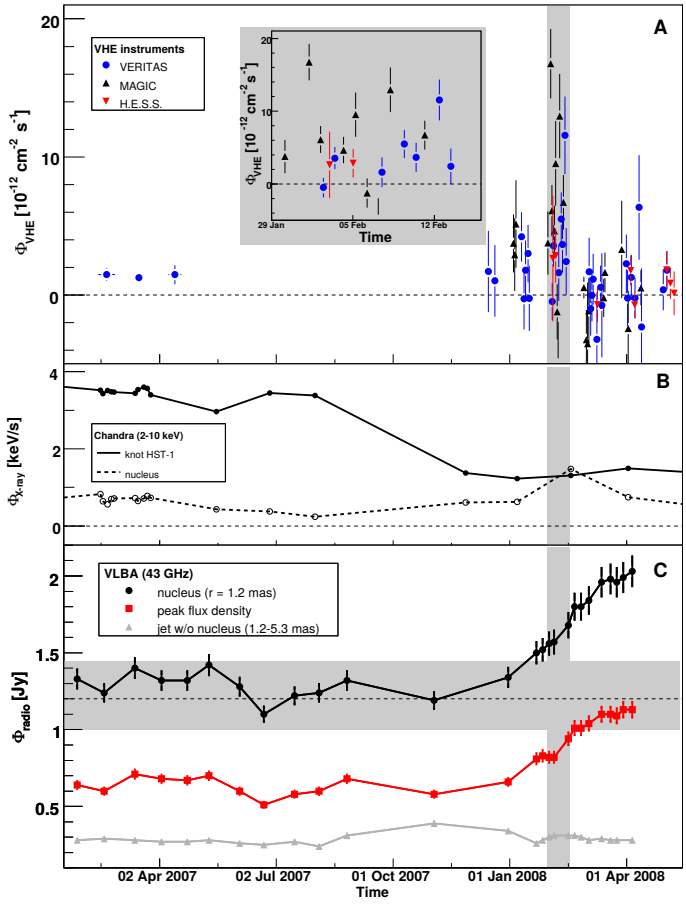


FIGURE 2.13: Combined M87 light curves in the VHE (top panel), X-ray (middle panel) and radio (bottom panel) ranges. The shaded horizontal area shows the typical radio flux before the flare. Image taken from (Acciari et al., 2009).

On the other hand, polarimetric observations by GRAVITY and the EHT strongly suggest the presence of dynamically important poloidal magnetic fields in the vicinity of the black hole. This supports the dominant role of magnetic fields in the launching of the jet. It seems that to produce sufficiently powerful jets, extraction of energy from the black hole is compulsory. As we already pointed out, this requires continuous injection of plasma into the magnetosphere. The origin of this plasma supply is still unknown. It is possibly related to the nature of the accretion flow (radiatively inefficient, geometrically thick, optically thin) which was confirmed by the EHT image of M87*.

This body of evidence points toward magnetospheric processes, occuring in the close environment of supermassive black holes, as the source of this activity.

Chapter 3

State of the art

In this section, I describe in more detail magnetospheric processes which can be responsible for VHE flares from radio galaxies, and describe the state of the art in numerical simulations of plasmas around black holes. Although I will adopt magnetospheric processes as a working hypothesis, it should be pointed out that alternative scenarios have been put forward to explain VHE flares, such as the jet-in-jet model (Giannios et al., 2009, 2010). Reviews of these mechanisms are provided by Rieger & Aharonian (2012), Aharonian et al. (2017) and Rieger & Levinson (2018) for example.

3.1 Magnetospheric gap models

3.1.1 Particle acceleration

As we will see in Sec. 6.1.3, spinning black holes threaded by poloidal magnetic fields in vacuum produce intense electric fields parallel to the magnetic field lines. In reality, the magnetosphere is most likely not vacuum, and the electric field could be screened by the plasma. The question amounts to whether the black-hole environment can produce plasma in sufficient amounts to screen the electric field everywhere, at all times. The minimum number density required to completely screen the electric field is the Goldreich-Julian density $n_{\rm GJ}$ (see Sec. 6.2 for more details). If the plasma density falls short of the Goldreich-Julian density, gaps can form in the magnetosphere where the electric field is not screened. This electric field can then accelerate particles to very high energies. Ultimately, it is the rotational energy of the black hole which is extracted that way. This mechanism had first been suggested in the context of pulsar magnetospheres (Sturrock, 1971; Ruderman & Sutherland, 1975). Vacuum breakdown in this context of this thesis was originally considered as a possible plasma source by Blandford & Znajek (1977).

For a black hole with spin parameter a, the maximum energy that a charged particle can be accelerated to is roughly given by $E_{\rm max} = qaB_0 r_{\rm g}$, if the particle exploits the full potential drop that has formed in a gap of characteristic size $r_{\rm g}$, with B_0 the typical magnetic strength. The latter can be estimated by assuming equipartition between the magnetic pressure ${B_0}^2/8\pi$ and the accreting gas pressure $\dot{m}c/r_{\rm g}^2$, which yields (Hirotani & Pu, 2016)

$$B_0 \simeq 10^4 \left(\frac{10^9\,{\rm M}_\odot}{M}\right)^{1/2} \left(\frac{\dot{m}}{\dot{m}_{\rm Edd}}\right)^{1/2} {\rm G}. \tag{3.1}$$

This yields lepton energies up to $10^{20}\,{\rm eV}$ (Levinson, 2000), making it possible to produce VHE flares by inverse Compton scattering for example (see the next section). Of course, the magnetic strength will be smaller if the black hole does not accrete at the Eddington rate, and the gap size can also be smaller than $r_{\rm g}$. In addition, the maximum energy can be limited by radiative losses, so $E_{\rm max}$ is a conservative upper limit.

3.1.2 Gap formation

The next question to address is whether spark gaps can actually form in the magnetosphere. Let us assume that the magnetosphere is initially filled with plasma. As we will see in Sec. 7.2.4, plasma close to the black hole must necessarily be sucked by the black hole, whereas an outflow develops at larger distance. Gaps will form perforce, unless plasma is perpetually replenished within the magnetosphere. Direct feeding of the polar regions of the magnetosphere by the accreting plasma is unlikely, because of the high magnetization at high latitudes, which pushes material to the equatorial plane, and the centrifugal force barrier. As a result, in the absence of an external plasma source, the polar funnel can become quickly depleted.

A possibility is plasma supply through photon-photon annihilation of thermal MeV photons emitted by the hot accretion flow (see Sec. 8.2). The resulting number density was evaluated by Levinson & Rieger (2011) and Hirotani & Pu (2016), who found that this density is lower than the Goldreich-Julian density for accretion rates below $\simeq 10^{-4}\dot{m}_{\mathrm{Edd}}$. This is a necessary condition for spark gaps to form, which can only be satisfied in low-luminosity AGN. As was shown in Sec. 2.2.1, this condition is readily verified in the case of M87*.

Since pair production by annihilation of MeV photons cannot screen the gap in low-luminosity AGN, these are able to accelerate particles. Levinson (2000) showed that unlike in the case of pulsar magnetospheres, here, vacuum breakdown does not occur by the emission of curvature photons, since that would require magnetic fields stronger than what is observed in AGN. Neronov & Aharonian (2007) also noted that gamma-ray emission from accelerated protons would be in tension with observations. On the other hand, they hypothesized that high-energy electrons or positrons could produce high-energy radiation by inverse Compton scattering of low-energy photons emitted by the accretion flow. They also checked that radiative losses of accelerated particles under the effect of curvature radiation and inverse Compton scattering still allowed a gamma-ray spectrum extending beyond $10\,\mathrm{TeV}$. Hirotani et al. (2016) found that the resulting gamma-ray fluxes could be observable for nearby AGN.

3.1.3 The importance of gaps

TeV photons interact preferentially with infra-red photons of energy $\simeq 1\,\mathrm{eV}$ (see Eq. (8.30)). Consequently, they can interact with eV photons emitted by the accretion flow, and trigger pair production by photon-photon annihilation. That way, the gamma-ray radiation can exert a backreaction on the gap by providing a plasma supply which could screen the electrostatic field. This question is of paramount importance in the context of black-hole magnetospheres. Blandford & Znajek (1977) realized that the magnetosphere had to be populated with plasma in order to extract energy from the rotating black hole. This necessity will be justified in detail in Sec. 6.2. They considered vacuum breakdown as a means to fill the magnetosphere with pair plasma. In order to efficiently extract energy from the black hole, there should be a continuous plasma source in the polar regions at least, in order to avoid complete depletion, as mentioned earlier. This is a necessary condition to activate the Blandford-Znajek process. Levinson & Rieger (2011) showed that this process could provide plasma in sufficient amounts so as to screen the electric field above the gap and activate the Blandford-Znajek mechanism. It follows from this picture that spark gaps could be a necessary feature of AGN models. They are not only invoked to explain VHE emission, but they also provide the crucial plasma supply which allows the black hole to power the jet.

The question of whether the produced gamma rays can come out of the emitting region finally comes up. Internal absorption of gamma rays is precisely due to pair creation, which could prevent them from leaving the magnetosphere. Let us assume that gamma rays and soft photons both originate from a spherical zone of radius R_{IR}, and let us denote L_{IR} the

luminosity in the IR band. The number density of target soft photons of energy ϵ is $n_{\rm IR} \simeq L_{\rm IR}/4\pi {R_{\rm IR}}^2 c\epsilon$. The cross section for pair production is of the order of $\sigma_{\rm T}$, so that the optical depth for gamma rays reads

$$\tau_{\gamma\gamma} \simeq n_{\rm IR}\sigma_{\rm T} R_{\rm IR} \simeq \frac{L_{\rm IR}\sigma_{\rm T}}{4\pi c R_{\rm IR}\epsilon}. \tag{3.2}$$

In the case of M87*, for an infra-red luminosity $L_{\rm IR} \simeq 10^{41}\,{\rm erg\cdot s^{-1}}$ (Whysong & Antonucci, 2004) and for TeV gamma rays, we have $\epsilon \simeq 1\,{\rm eV}$ and $\tau_{\gamma\gamma} \simeq 10^2\,(r_{\rm g}/R_{\rm IR})$ (Neronov & Aharonian, 2007). The size of the infra-red source on M87* is hard to constrain, but it is unlikely to be much smaller than $10^2 r_{\rm g}$. This would imply that gamma rays are actually able to escape from the magnetosphere. This conclusion can be generalized to supermassive black holes accreting in a radiatively inefficient way, for which $L_{\rm IR}$ is much smaller than $L_{\rm Edd}$ (Wang et al., 2008; Levinson & Rieger, 2011; Rieger & Aharonian, 2012). However, it is very sensitive to the accretion rate and the spin of the black hole, or the exact model for the RIAF. On the opposite, gamma rays could not escape if the black hole accreted through a standard thin disk (Brodatzki et al., 2011): this might explain why VHE emission is a feature of low-luminosity radio-galaxies.

3.1.4 Gap dynamics

The first gap models assumed stationarity. The gap was taken to be a small localized perturbation of the magnetosphere, where the electric field was not screened. Let us suppose that the magnetosphere rotates at an angular velocity Ω and that the magnetic field is close to monopolar. In the laboratory frame, the rotationally induced electric field is orthoradial everywhere, but a radial component may arise in the gap. For clarity, we neglect general relativistic corrections. Noting $\boldsymbol{E}'$ the electric field measured in the frame rotating with magnetosphere, Maxwell-Gauss' law reads (Rieger & Levinson, 2018)

$$\boldsymbol{\nabla}\cdot\boldsymbol{E}' = 4\pi\left(\rho - \rho_{\rm GJ}\right), \tag{3.3}$$

where ρ is the electric charge density and $\rho_{\rm GJ} = -\boldsymbol{\Omega}\cdot\boldsymbol{B}/2\pi c$ is the Goldreich-Julian density (see Sec. 6.2). No radial electric field is induced if $\rho = \rho_{\rm GJ}$. The Goldreich-Julian density depends on the magnetic geometry. At the inner and outer gap boundaries, the electric field should be screened, so that $\boldsymbol{E}'$ vanishes. This implies that the radial component E'_r has a maximum or a minimum inside the gap, and that $\rho - \rho_{\rm GJ}$ has opposite signs at the two boundaries (Katsoulakos & Rieger, 2018).

As it happens, in a spinning black-hole magnetosphere, even if the magnetic field is radial everywhere, there can be a "null surface" where $\rho_{\rm GJ} = 0$ (it is shown in red in Fig. 3.1(a)). This is a purely general relativistic effect. It has been argued that because the gap is charge-starved, a change of sign of $\rho - \rho_{\rm GJ}$ happens preferentially at a null surface (Beskin et al., 1992; Hirotani & Okamoto, 1998; Hirotani & Pu, 2016). The behavior of these quantities is shown in Fig. 3.1(b). A similar line of reasoning had been followed in the context of pulsar gap emission (Cheng et al., 2000), the existence of a null surface being due to the dipolar magnetic geometry this time. This location was especially appealing, as it allowed for the construction of stationary gap models. The radiative properties and structure of the gap have been studied analytically and numerically with this preliminary assumption (Neronov & Aharonian, 2007; Vincent & LeBohec, 2010; Ptitsyna & Neronov, 2016; Katsoulakos & Rieger, 2018, 2020). The emitted spectrum includes a component from curvature emission around $\simeq 1\,{\rm GeV}$ and an inverse Compton component centered on $\simeq 10\,{\rm TeV}$. Alternatively, the MHD stagnation surface separating inflow and outflow in single-fluid

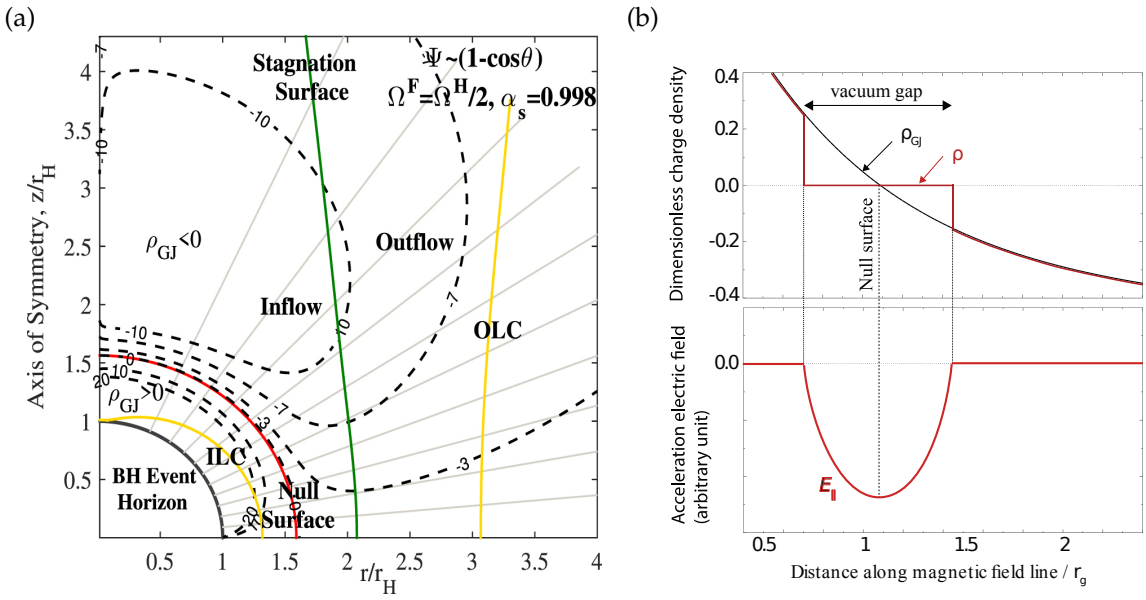


FIGURE 3.1: (a) Sketch of a black-hole magnetosphere in the poloidal plane, in a monopole magnetic configuration (solid gray lines) and with a spin parameter $a = 0.998$. The yellow lines indicate the inner and outer light surfaces. The dashed lines are contour levels of $\rho_{\rm GJ}$, with the red line showing the null surface ($\rho_{\rm GJ} = 0$). The MHD stagnation surface is shown in green. Figure taken from Katsoulakos & Rieger (2020). (b) Distribution of charge density (red solid line, upper panel) and parallel electric field (lower panel) along a magnetic field line, in the case of a gap located at the null surface. The Goldreich-Julian density is also shown in the upper panel (black solid line). Figure taken from Hirotani & Pu (2016).

MHD (Takahashi et al., 1990) has been suggested as a plausible position for the gap (Broderick & Tchekhovskoy, 2015). This surface corresponds to a force balance between the gravitational, centrifugal and electromagnetic stresses. Its location depends on the plasma loading of the field lines (Globus & Levinson, 2013). This surface is also shown in green in Fig. 3.1(a).

In contrast, Levinson & Segev (2017) argued against stationary gaps. They found that steady-state solutions are inconsistent with the presence of a stagnation surface. Besides, stationarity implies that every accelerated lepton only creates one pair (the "closure" condition). This places an upper bound on the accretion rate which is too low to match observational constraints. Hence, spark gaps are likely to be intermittent, setting them out of reach of analytical computations. In addition to that, Levinson et al. (2005) studied the time-dependent dynamics of spark gaps, and showed that the response of the electric field to a burst of pair creation was oscillatory in nature, due to a time-varying electric current. They suggested that steady-state gaps were unstable to temporal perturbations. This work had put the emphasis on pulsar magnetospheres, but it represents fairly well the gap oscillatory dynamics, as we will see in Chap. 12.

3.2 Numerical modeling

3.2.1 The need for simulations

In the previous section, we have discussed how the local gaps can have a global impact on the magnetosphere, as they control the plasma supply. Conversely, the global magnetospheric structure affects the gap. It is crucial to determine this structure in order to interpret the multi-wavelength emission from AGN and to understand how jets are formed. The

natural framework to study black-hole magnetospheres as a whole has long been the force-free approximation, which consists in neglecting the plasma inertia. This is the approach taken by Blandford & Znajek (1977) and MacDonald & Thorne (1982), which is described in Chap. 6 and 7. Unfortunately, only a restricted number of solutions can be found this way. Even the exact solution from Blandford & Znajek (1977) is only valid at low spin parameter. What is worse, it is not clear whether the force-free approximation is valid in the first place. For example, it was challenged by Punsly & Coroniti (1990a,b) and Punsly (2001), who claimed that particle inertia could not be overlooked. They also questioned the role of the event horizon in the Blandford-Znajek, as well as the long-term stability of this solution (more details can be found in Chap. 7). Instead, they proposed a model in which particle inertia was a key ingredient. Besides, as will be shown in 6.2, the force-free appoximation must itself break down in a current sheet, casting doubt on the relevance of this approach.

Numerical simulations of a black-hole magnetosphere became possible in the 2000's, with the increase in computational power. A first step was taken by Komissarov (2001) and Komissarov & McKinney (2007) who showed by the means of GR force-free simulations that the Blandford-Znajek solution was actually asymptotically stable. The structure of the a force-free simulation is shown in Fig. 3.2(a). This confirmed that the Blandford-Znajek solution was valid, but the question of the importance of particle inertia remained. In particular, force-free codes cannot handle dissipative current sheets well.

3.2.2 MHD simulations

Ideal GRMHD techniques have been developed to simulate the close environment of black holes (e.g., Gammie et al., 2003). Komissarov (2004b) performed GRMHD simulations of a monopole magnetosphere and retrieved the results of GR force-free simulations to an excellent degree (see Fig. 3.2(b)). Koide (2003) and Komissarov (2005) have simulated black holes in an initially vertical magnetic field, and find that a stable equatorial current sheet

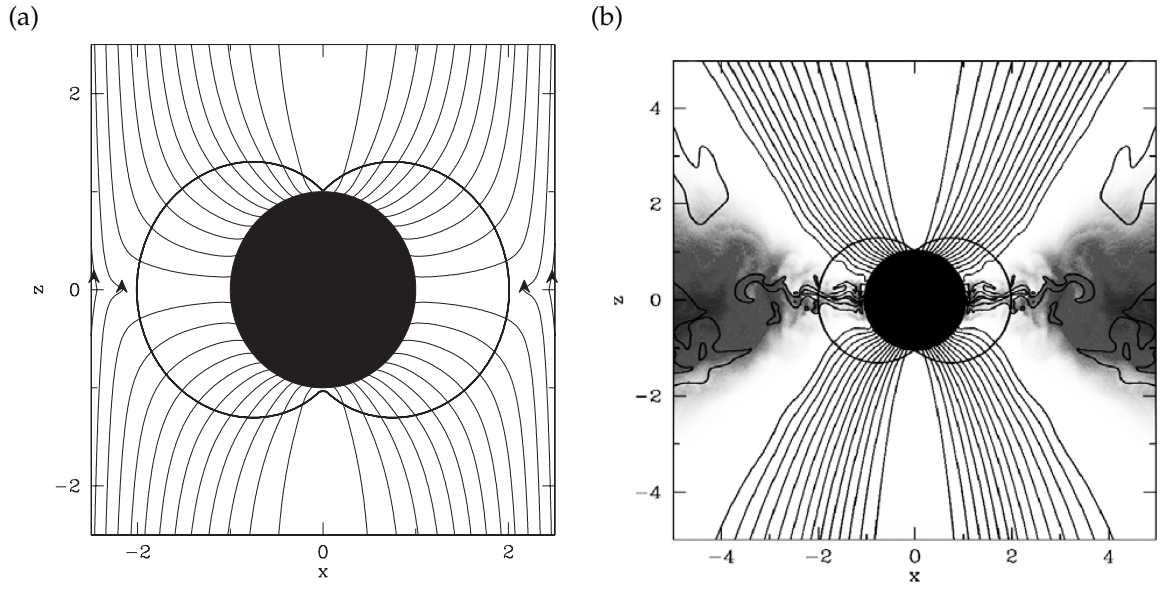


FIGURE 3.2: (a) Steady-state force-free simulation of a black hole immersed in a vertical magnetic field. The magnetic flux is pulled onto the event horizon. (b) Quasi-steady state of a GRMHD simulation of an accretion flow around a Kerr black hole with $a = 0.999$. The solid lines show poloidal magnetic field lines. Figures taken from Komissarov & McKinney (2007).

naturally develops. GRMHD simulations of the close magnetosphere seem to validate the Blandford-Znajek process and confirm the validity of the force-free framework, without any imperious need for particle inertia.

Such simulations have then been used to simulate accretion flows more realistically (De Villiers et al., 2003; McKinney & Gammie, 2004). The underlying assumptions of these simulations are that radiation can be neglected, and that the plasma can be treated as a fluid rather than as a collisionless plasma. In accreting low-luminosity AGN, the dynamical effect of radiation is subdominant by definition, so that one can afford to discard radiation from the simulation. This kind of accretion flows is easier to simulate than optically thick, geometrically thin accretion disks. Most of these simulations start with a magnetized thick torus in equilibrium, which is destabilized by the magnetorotational instability (MRI). The accretion flow is optically thin and geometrically thick. Matter then accretes toward the black hole, transporting magnetic flux to the event horizon. These simulations demonstrated clearly the activation of the Blandford-Znajek mechanism (Tchekhovskoy et al., 2011), and the launching of large-scale Poynting-dominated jets (McKinney, 2006).

Ideal GRMHD simulations are scale-free. One of the most important dimensionless parameters is the magnetic flux threading the black hole Φ when the simulation reaches a steady state. The magnetic energy density becomes comparable with the accretion energy density when $B^2/8\pi \simeq \left(\dot{m}/4\pi r_{\rm g}^2 c\right) c^2$. Thus, the normalized magnetic flux ϕ should be defined as

$$\phi = \frac{\Phi}{\sqrt{\dot{m} r_{\rm g}^2 c}}. \tag{3.4}$$

The final magnetic flux ϕ is controlled by the initial magnetization of the torus. If the magnetic field supplied to the black hole is too strong, ϕ is found to saturate at values $\simeq 50$. The MRI has little effect on the dynamics of the flow. This is often called the "magnetically arrested disk" (MAD) configuration. In this case, the magnetic field is dynamically important and impacts the dynamics of the inflow (Narayan et al., 2003; Tchekhovskoy et al., 2011; McKinney et al., 2012). Recently, the MAD state has been observed to arise from a a new set of initial conditions, drawn from larger-scale MHD simulations (Ressler et al., 2020), or from a non-rotating flow with an initially uniform magnetic field (Ressler et al., 2021). This suggests that the MAD configuration can develop quite naturally in more general setups. All in all, the gas reaching the event horizon carries a small amount of angular momentum but a large amount of magnetic flux. Otherwise, if ϕ does not saturate and remains close to 1, the disk does not become magnetically arrested, in what is called the "standard and normal evolution" (SANE) model. In such accretion flows, the MRI is very active. Both models have been extensively used to produce synthetic images of black-hole magnetospheres and interpret the EHT data (Event Horizon Telescope Collaboration et al., 2019b).

On the other hand, the second core assumption of GRMHD simulations is likely to fail in low-luminosity AGN. The mean free path l_e for electron-electron Coulomb collisions is given in terms of the plasma parameter Λ and the plasma frequency $\omega_{\rm p}$ by (Fitzpatrick, 2015)

$$l_e \simeq \frac{\Lambda}{\ln(\Lambda)} \frac{c}{\omega_{\rm p}}; \quad \Lambda = \frac{(k_{\rm B} T)^{3/2}}{e^3 \sqrt{n}}; \quad \omega_{\rm p} = \sqrt{\frac{4\pi n e^2}{m_{\rm e}}}, \tag{3.5}$$

with T the temperature and n the density of the plasma. Numerical application in the case of M87*, with $T \simeq 10^{10}\,{\rm K}$ and $n \simeq 10^4\,{\rm cm}^{-3}$, yields $\ln(\Lambda) \simeq 40$, so that $l_e/r_{\rm g} \simeq 10^5 \gg 1$. Because the accretion rate is so low, the plasma is so hot and tenuous that it is collisionless, and the electron or proton mean free path is much larger than the typical size of the system. The situation might be less dire as it seems, as other physical processes (such as wave-particle

interactions; see Kunz et al., 2014) can isotropize the system. Promising efforts have been made to include nonideal effects and radiative cooling into GRMHD simulations (Foucart et al., 2016; Sądowski et al., 2017; Ryan et al., 2018). Nonetheless, the fluid description is not adequate for dealing with hot accretion flows.

The MHD approach suffers from several other shortcomings. The funnel region, close to the polar axis of the black hole, is predicted to have very small densities compared to the accretion flow. This is because the angular momentum of matter is generally small in this region, and plasma is either sucked in or ejected in an outflow. Besides, the funnel is generally highly magnetized (Porth et al., 2019), so that plasma from the accretion flow cannot penetrate it, and its density and pressure can be very low. This is very challenging for MHD numerical schemes, which are often forced to employ density floors to prevent vacuum from forming (McKinney & Gammie, 2004; Porth et al., 2019). Hence, although there have been attempts to evaluate pair production from radiative GRMHD simulations in post-processing (Mościbrodzka et al., 2011; Wong et al., 2021), these simulations cannot model self-consistently the plasma supply in the jet region, which is probably induced by the presence of spark gaps in the inner parts of the magnetosphere. Besides, fluid simulations cannot constrain particle acceleration and nonthermal radiation, because the dynamics of individual particles are not followed. Because the microphysics of the plasma is not included, dissipative regions (such as the equatorial current sheet, see Chap. 13) are often governed by numerical resistivity (however, see Ripperda et al. 2020 for recent resistive GRMHD simulations).

3.2.3 Kinetic simulations

Force-free analytical approaches and GRMHD simulations teach us a great deal about the global dynamics of the system, and the large-scale magnetic configuration. Yet, in order to interpret the nonthermal emission from AGN, it seems necessary to go beyond and treat the problem from first principles. Particle-in-cell (PIC) methods are *ab initio* kinetic plasma simulations, which can include self-consistently a variety of physical phenomena. They will be described in detail in Chap. 9. Several crucial steps were taken very recently in the modeling of this problem, in order to go beyond standard plasma simulations. First, radiative transfer processes must be implemented. This had already been done in 1D simulations of pulsar polar cap discharges (Timokhin, 2010; Timokhin & Arons, 2013; Cruz et al., 2021), including curvature emission and pair production by magnetic conversion. Chen et al. (2018) performed 1D simulations in flat spacetime with the adequate radiative processes: inverse Compton scattering and pair production by photon-photon annihilation.

Then, general relativity must be included. Levinson & Cerutti (2018) used the first general relativistic PIC simulations in 1D, with inverse Compton scattering and pair production to study the discharge of an initially empty gap. They treated the gap as a small perturbation of a global magnetosphere, and found that the gap evolved to a quasi-steady state where the parallel electric field was kept at a very small amplitude and varied randomly. Chen & Yuan (2020) performed similar simulations, with $\rho = \rho_{\rm GJ}$ initially everywhere in the gap, but arrived at a different conclusion. They found a cyclic behavior, where a gap periodically opened at the null surface. Kisaka et al. (2020) extended the range of parameters and initial conditions investigated by Levinson & Cerutti (2018) using the same code, with much longer integration times. In the case of a global current consistent with the large-scale magnetospheric configuration, they also found that the electric field opened periodically at the null surface, although the resulting plasma density seemed stationary. All these 1D simulations focused on a single radial field line in a monopolar magnetic configuration. Note that 1D simulations are purely electrostatic, and only evolve the radial component of the electric field. The gap is assumed to be embedded in a force-free, stationary magnetosphere, which

imposes external conditions on the gap. It is considered as a small perturbation, which controls plasma supply without affecting its global properties. For example, in 1D GRPIC simulations, the current flowing through the gap and the angle of the single simulated field line are free parameters.

These simulations are able to capture the dynamics of the gaps, but must be educated with the global structure. In parallel, Parfrey et al. (2019) presented the first global 2D, fully general relativistic PIC code. The foundations of their code are described in Chap. 10. They simulated a black hole in a vertical magnetic field, with the aim of retrieving a nearly force-free magnetosphere. To that end, they used a particular plasma injection method, similar to that used by Belyaev (2015b). They evaluated the local magnetization and the normalized unscreened electric field $\boldsymbol{D} \cdot \boldsymbol{B}/\boldsymbol{B}^2$. Wherever the unscreened electric field was above a critical value $\varepsilon_{\rm th}$ (typically between 10^{-2} and 10^{-3}) and the magnetization was not too low, plasma was injected with a density

$$\delta n \simeq \frac{1}{4\pi e r_{\rm g}} \frac{\boldsymbol{D} \cdot \boldsymbol{B}}{B}. \tag{3.6}$$

This choice of injection method allowed them to mimic pair creation by vacuum breakdown, when the unscreened electric field became too large. Their results are shown in Fig. 3.3. In addition to reproducing aspects of force-free magnetospheres, they were able to model features only accessible to kinetic simulations: the reconnecting current sheet and the impact

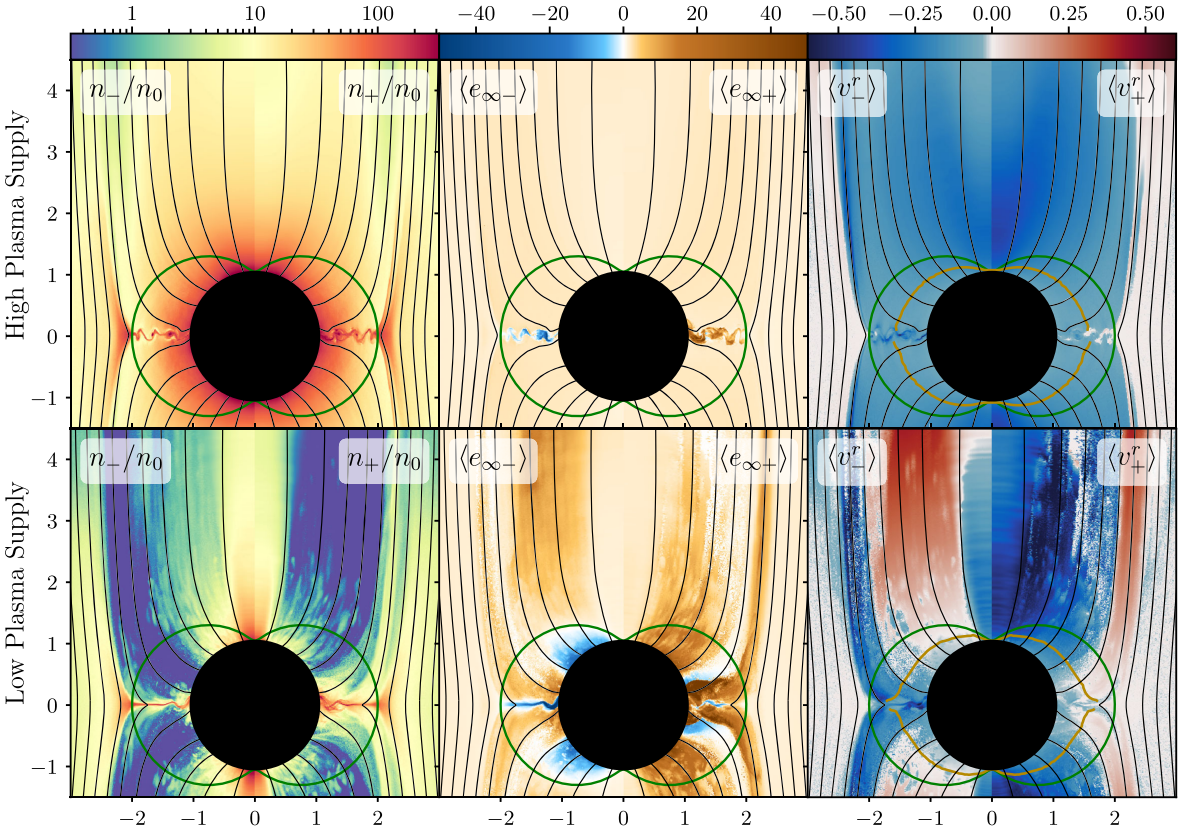


FIGURE 3.3: Steady-state density, energy-at-infinity and radial 3-velocity for electrons (left) and positrons (right) from the global general relativistic kinetic simulations by Parfrey et al. (2019). The upper panel shows a simulation with $\varepsilon_{\rm th} = 10^{-3}$, whereas the lower panel has $\varepsilon_{\rm th} = 10^{-2}$. All distances are in units of $r_{\rm g}$. Figure taken from Parfrey et al. (2019).

of negative-energy particles, for instance. Importantly, they showed how the solution depended on the physical parameter $\varepsilon_{\rm th}$, which is in principle governed by radiative transfer physics. In the case of a high threshold $\varepsilon_{\rm th}$, the funnel became charge-starved, and pair creation occured intermittently. However, their simplified choice of injection method precluded any chance of seeing a gap develop, as plasma supply was not treated self-consistently.

3.3 Objectives of this thesis

Plasmas around AGN are collisionless and optically thin. Their dynamics and nonthermal emission cannot be well captured by fluid simulations, and instead require a kinetic treatment, which can account for the presence of gaps, in particular. Unlike previous simulations of magnetospheric gaps, I aim for global simulations, relaxing the need for prescriptions encoding the external environment. From the state-of-the-art kinetic simulations described in Sec. 3.2.3, my goal is to take the next logical step, and implement global, radiative, and general relativistic 2D PIC simulations of black-hole magnetospheres. In order to take full advantage of these simulations in interpreting observations of AGN, I also devise a method to obtain synthetic observables from my simulations, by coupling the PIC code to a ray-tracing code and tracking the emitted photons from the magnetosphere to the observer. Probably too ambitiously, the ultimate objective of this thesis would be to address some of the following questions, which are currently out of reach of numerical simulations.

1. Can the Blandford-Znajek mechanism be activated in simulations including *ab initio* plasma supply?
2. Where do gaps and particle acceleration take place in the magnetosphere?
3. What is the radiative signature of black-hole magnetospheres in the very high-energy range?
4. Can the EHT and GRAVITY observations be produced by pure magnetospheric processes?

This thesis is outlined as follows. The motivations, observational context and numerical state of the art relevant to this thesis have been presented in this part (Part I). Part II deals with the theoretical foundations of this thesis and the technicalities. The $3+1$ formalism of general relativity, black-hole electrodynamics, its application to the Blandford-Znajek mechanism, and ubiquitous radiative processes in active galactic nuclei are outlined. Part III contains a description of the numerical techniques used in this thesis. The now well established PIC methods are summarized. The general relativistic specificities are described in detail. Although I did not design the `GRZeltron` version in the first place, these details and derivations have not been published yet. I also describe the radiative transfer algorithm that I designed for the study of spark gaps. In Part IV, I present the results of my simulations, along with the observational consequences. These results have mostly been published in Crinquand et al. (2020) and Crinquand et al. (2021). A summary of my findings and some perspectives are given in Part V.

Part II

Foundations

"Le principe consista d'abord dans votre conviction que nous n'atteindrons jamais le fond des choses, non en vertu d'une malédiction ou de la faiblesse de nos facultés, mais pour la raison définitive et radicale que les choses n'ont pas de fond."

"The principle first consisted in your belief that we will never get to the bottom of things, not for some curse or the weakness of our abilities, but for the definitive and radical reason that things have no bottom."

JÉRÔME FERRARI, Le principe

Chapter 4

Kerr black holes

According to the no-hair theorem, black holes are characterized by a very limited number of parameters: their mass, electric charge and angular momentum. It is generally believed that astrophysical black holes carry a negligible charge, in the sense that the characteristic associated length scale $r_{\mathrm{Q}} = Q^2 G/4\pi c^4$ is small with respect to the graviational radius $r_{\mathrm{g}} = GM/c^2$, for a black hole of mass M and charge Q. This is because significantly charged black holes would be able to accrete a preferred sign of charges, until it gets neutralized. Angular momentum, on the other hand, cannot be so easily removed.

Studying spinning black holes not only has theoretical foundations, but also observational motivations. Most stellar-mass black holes are the remnants of massive stars. Thus, they are expected to inherit some of the angular momentum of the primary star. Even if a black hole is originally formed without angular momentum, it can spin up during its subsequent evolution, as a result of accretion, or after coalescence with another compact object. Unfortunately, measuring the spin of a black hole is a huge endeavour, and there exists only a few trustworthy measurements. In this field, theoretical modeling precedes compelling observations, but great progress has been made. Different measurements techniques are described in the reviews by Reynolds (2013, 2020).

Spinning black holes are oft-referred to as "Kerr" black holes, in honor of the mathematician Roy Kerr, who was the first to derive an exact solution for black holes possessing angular momentum. Black-hole spin can actually have a significant effect on astrophysical phenomena, the most striking example being the possibility to produce a jet by tapping the black-hole rotational energy. Thereby, the title of this thesis is a testament to the importance of Kerr black holes and the need to understand their specificities.

The basics of general relativity covered in this section were mostly drawn from the textbooks by Frolov & Zelnikov (2011) and Carroll (2003), along with the most excellent online lecture notes by Blau (2018). The reader is referred to these sources for further insight.

4.1 Spinning black holes

4.1.1 The many facets of the Schwarzschild event horizon

A black hole can be loosely defined as a "localized region of spacetime from which neither massive particles nor massless ones can escape" (Gourgoulhon, 2021). This object must possess a boundary, the event horizon, which no particle initially inside it can cross: it is a one-way membrane. Besides, its mass must lie entirely within the event horizon. The simplest black hole is described by the Schwarzschild solution. However, the definition of the event horizon of a Schwarschild black hole can be ambiguous. Our goal is to compare the special surfaces of a Kerr black hole compared to a Schwarzschild black hole, in order to better understand the specificities of Kerr black holes.

Let M be the mass of a Schwarzschild black hole, and $r_{\mathrm{s}} = 2GM/c^2 = 2r_{\mathrm{g}}$ be the Schwarzschild radius. The metric $g_{\mu\nu}$ of the spacetime reads, in the standard Schwarzschild

coordinates (t, r, θ, φ) and with a $(-,+,+,+)$ signature:

$$\mathrm{d}s^2 = -f(r)c^2\,\mathrm{d}t^2 + f(r)^{-1}\,\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right), \tag{4.1}$$

where $f(r) = 1 - 2r_\mathrm{g}/r$. Two things of interest happen at $r = r_s$, which is the only root of the function f:

1. $g_{tt}(r_\mathrm{s}) = f(r_\mathrm{s}) = 0$;
2. $g^{rr}(r_\mathrm{s}) = 1/g_{rr}(r_\mathrm{s}) = f(r_\mathrm{s}) = 0$.

Condition 1, which is *a priori* coordinate-dependent, can be interpreted in the following way. Static observers at fixed (r, θ, φ), which have a 4-velocity $\boldsymbol{u} \propto \partial_t$, can only exist for $r > r_\mathrm{s}$. Indeed, the 4-vector ∂_t becomes null if $g_{tt} = 0$, and remains timelike for $r > r_\mathrm{s}$. The hypersurface $\mathcal{H}$ defined by $r = r_\mathrm{s}$ is the limit surface for static observers. Incidentally, it is also an infinite redshift surface. Condition 2 means that the vector $\boldsymbol{N} = \partial_r$, which is normal to surfaces of constant r, is null at $r = r_\mathrm{s}$ and becomes timelike for $r < r_\mathrm{s}$. $\mathcal{H}$ is a null hypersurface, which can therefore only be crossed once by a timelike wordline.

In a Schwarzschild black hole, these two properties of the surface $\mathcal{H}$ coincide. This would also be true if the black hole had a significant electric charge; however we will see that they differ for a spinning black hole.

4.1.2 The Kerr metric

This metric describes a stationary (but not static) spacetime. In the Boyer-Lindquist coordinates (t, r, θ, φ), the line element of the Kerr metric reads

$$\mathrm{d}s^2 = -\left(1 - \frac{2r_\mathrm{g}r}{\Sigma}\right)c^2\,\mathrm{d}t^2 - \frac{4ar_\mathrm{g}^2 r\sin^2\theta}{\Sigma}c\,\mathrm{d}t\,\mathrm{d}\varphi + \frac{\Sigma}{\Delta}\,\mathrm{d}r^2 + \Sigma\,\mathrm{d}\theta^2 + \frac{A\sin^2\theta}{\Sigma}\,\mathrm{d}\varphi^2, \tag{4.2}$$

where

$$\Sigma = r^2 + a^2r_\mathrm{g}^2\cos^2\theta\,; \quad \Delta = r^2 - 2r_\mathrm{g}r + a^2r_\mathrm{g}^2\,; \quad A = \left(r^2 + a^2r_\mathrm{g}^2\right)^2 - a^2r_\mathrm{g}^2\Delta\sin^2\theta. \tag{4.3}$$

The Boyer-Lindquist coordinates reduce to spherical coordinates as $r \to +\infty$. The Kerr metric depends on two parameters only: the gravitational radius r_g and the spin parameter a. It reduces to the Schwarzschild metric when $a \to 0$. The presence of a non-zero $g_{t\varphi}$ breaks the symmetry by reflection of time $t \to -t$, but retains the symmetry $(t \to -t\,;\ \varphi \to -\varphi)$. The physical meaning of a can be understood by looking at the weak-field limit. For $r \gg r_\mathrm{g}$, the metric expand to lowest order in r_g/r as

$$\mathrm{d}s^2 \simeq -\left(1 - \frac{2r_\mathrm{g}r}{\Sigma}\right)c^2\,\mathrm{d}t^2 - \frac{4ar_\mathrm{g}^2 r\sin^2\theta}{\Sigma}c\,\mathrm{d}t\,\mathrm{d}\varphi + \left(1 + \frac{2r_\mathrm{g}r}{\Sigma}\right)\left(\mathrm{d}r^2 + r^2\,\mathrm{d}\theta^2 + r^2\sin^2\theta\,\mathrm{d}\varphi^2\right). \tag{4.4}$$

This is precisely the metric of the spacetime surrounding a slowly rotating massive sphere, of mass M and angular momentum J, where the spin parameter is related to J by

$$a = \frac{Jc}{GM^2}. \tag{4.5}$$

Hence, one can interpret the spin parameter a that characterizes the Kerr metric as the normalized angular momentum of the black hole. It is complex to define properly the angular momentum of a black hole (see Sec. 6.1.3).

The Kerr metric is one of the most important solutions of Einstein's equations, for astrophysical purposes. This comes from several uniqueness theorems of paramount importance. Provided the black hole is uncharged, and some regularity assumptions are made, one can show the following theorems (see Chruściel et al. 2012 for an extended discussion).

- Any static vacuum solution of Einstein's equations in an asymptotically flat spacetime is spherically symmetric and coincides with the Schwarzschild metric (Israel, 1967).
- Any stationary and axially symmetric vacuum solution of Einstein's equations in an asymptotically flat spacetime is described by the Kerr metric. In particular, it is uniquely described by the values of two parameters (mass and angular momentum) (Carter, 1971; Robinson, 1975).

The Kerr metric can be generalized into the Kerr-Newman metric, to account for the presence of electric and magnetic charges. The solution can be found in Frolov & Novikov (1998). We also give the expression of the coefficients of the inverse metric in BL coordinates:

$$g^{00} = -\frac{A}{\Delta\Sigma}, \quad g^{rr} = \frac{\Delta}{\Sigma}, \quad g^{\theta\theta} = \frac{1}{\Sigma}, \quad g^{0\varphi} = -\frac{2ar_{\mathrm{g}}^2 r}{\Delta\Sigma}, \quad g^{\varphi\varphi} = \frac{\Delta - a^2 r_{\mathrm{g}}^2 \sin^2\theta}{\Delta\Sigma\sin^2\theta}. \tag{4.6}$$

4.1.3 Ergosphere

Just like in Sec. 4.1.1, the limit of static observers can be found by determining when ∂_t becomes null, that is, by determining the zeros of g_{tt}. We find that

$$g_{tt}(r,\theta) = 0 \iff \Sigma - 2r_{\mathrm{g}} r = r^2 - 2r_{\mathrm{g}} r + a^2 r_{\mathrm{g}}^2 \cos^2\theta = 0, \tag{4.7}$$

so that the limit of static observers is defined by

$$r_{\mathrm{e}}(\theta) = r_{\mathrm{g}}\left(1 + \sqrt{1 - a^2\cos^2\theta}\right). \tag{4.8}$$

The surface $\mathcal{S}$ defined by Eq. (4.8) is called the *ergosphere*, which borders the ergoregion (see the solid line in Fig. 4.1).

> This surface was named after the ancient Greek word *ergon*, meaning "work". This denomination is related to the special role that the ergosphere plays in the Penrose process (see Sec. 4.2.2).

Since the metric is invariant by translations of both t and φ (it possesses two Killing vectors $\boldsymbol{\eta} = \partial_t$ and $\boldsymbol{\psi} = \partial_\varphi$, see Sec. 4.2.1), one can define a more general class of stationary observers who remain at constant (r,θ), but rotate in the φ direction with constant angular velocity Ω. Their 4-velocity can be expressed as

$$\boldsymbol{u} = u^t\left(\boldsymbol{\eta} + \frac{\Omega\,(r,\theta)}{c}\boldsymbol{\psi}\right). \tag{4.9}$$

To be a little more precise, Ω gives the rate of variation of φ with respect to t, but not with respect to the proper time of the observer τ. To see this, one must recall from the definition of the 4-velocity that $u^\varphi = \mathrm{d}\varphi/\,\mathrm{d}\tau$ and $u^t = \mathrm{d}t/\,\mathrm{d}\tau$. Along the wordline of the observer, we have $\mathrm{d}\varphi/\,\mathrm{d}t = u^\varphi/u^t = \Omega$. Ω can also be interpreted as the angular velocity of the observer measured by a static observer at infinity, because its proper time matches t.

While static observers cannot exist within the ergoregion, stationary ones could. Such observers can exist if $\boldsymbol{\eta} + (\Omega(r,\theta)/c)\,\boldsymbol{\psi}$ is timelike, which translates into the following condition:

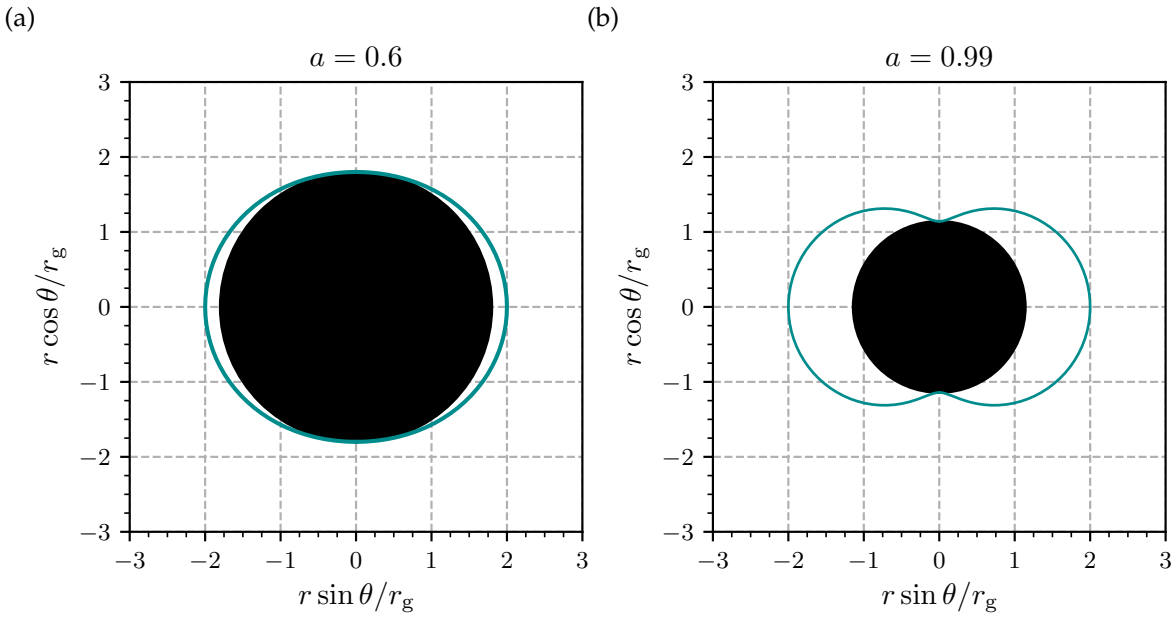


FIGURE 4.1: Poloidal cut of a the event horizon (black disk) and ergosphere (blue solid line) for $a = 0.6$ (a) and $a = 0.99$ (b).

$$f(\Omega, r, \theta) = g_{00} + 2\frac{\Omega}{c}g_{0\varphi} + \frac{\Omega^2}{c^2}g_{\varphi\varphi} < 0. \tag{4.10}$$

The solution of this quadratic inequality in Ω leads to

$$\Omega_-(r,\theta) < \Omega < \Omega_+(r,\theta), \tag{4.11}$$

where

$$\Omega_\pm = c\frac{-g_{0\varphi} \pm \sqrt{g_{0\varphi}^2 - g_{00}g_{\varphi\varphi}}}{g_{\varphi\varphi}}. \tag{4.12}$$

The situation is pictured in the (r, Ω) diagram of Fig. 4.2, for $\theta = \pi/2$, which shows which positions and angular velocities are permitted for a stationary observer. On the ergosphere $\mathcal{S}$, we have $g_{tt} = 0$ by definition, so that $\Omega_- = 0$: Ω_- is negative outside the ergosphere, and positive inside. This means that stationary observers can exist even within the ergosphere, but to do so they must rotate with the black hole ($\Omega_- > 0$), whereas outside the ergosphere they can also rotate against it. In that sense, it is often said that spinning black holes "drag" spacetime around them. This notion, sometimes called the Lense-Thirring effect, will be crucial to get an intuitive understanding of black-hole electrodynamics.

Besides, we note that the ergosphere is not a null surface. Actually, it is a timelike surface, and it cannot be a one-way membrane. Its normal vector has covariant components $n_\mu = (0, 1, -\mathrm{d}r_\mathrm{e}/\mathrm{d}\theta, 0)$, so that is norm is $n_\mu n^\mu = g^{rr} + g^{\theta\theta}(\mathrm{d}r_\mathrm{e}/\mathrm{d}\theta)^2 = \left(\Delta + (\mathrm{d}r_\mathrm{e}/\mathrm{d}\theta)^2\right)/\Sigma \geq 0$: $\boldsymbol{n}$ is spacelike. All in all, the ergosphere is very different from the definition that we gave of an event horizon.

> The condition $\Omega_- > 0$ continues to hold even if the observer has a momentum in the r or θ direction (Frolov & Zelnikov, 2011). Particles inside the ergosphere can move both at increasing or decreasing r as long as $\Omega > \Omega_-$.

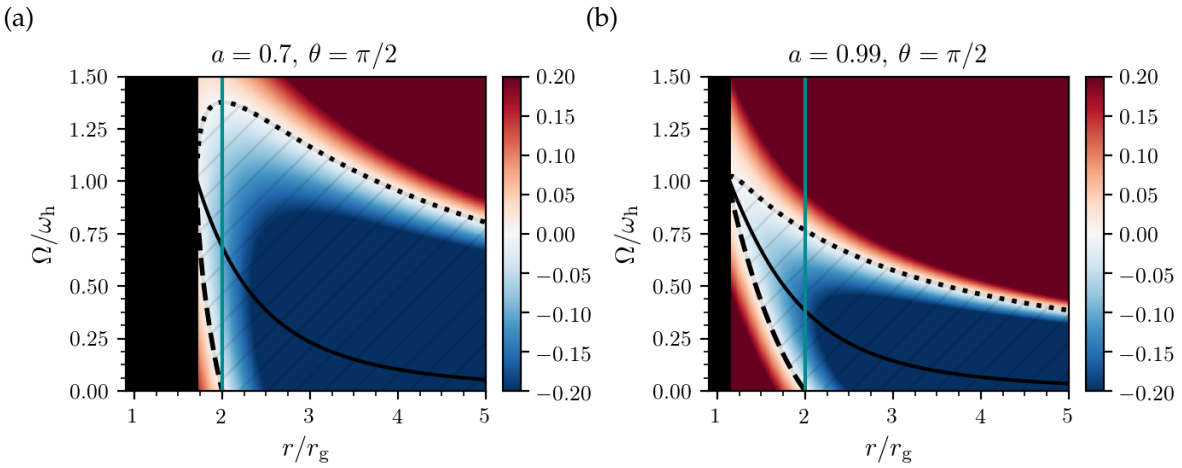


FIGURE 4.2: Map of the light surface function f (Eq. (4.10)) in the (r, Ω) plane, at a given angle $\theta = \pi/2$ for $a = 0.7$ (a) and $a = 0.99$ (b). In the blue hatched region, rotation at angular velocity Ω is subluminal ($f < 0$), whereas red color means that it is superluminal ($f > 0$). The black region indicates the black hole interior, whereas the vertical line marks the position of the ergosphere at $\theta = \pi/2$. The dotted line shows the evolution of $\Omega_+(r, \pi/2)$, whereas the dashed line marks $\Omega_-(r, \pi/2)$.

4.1.4 The Kerr event horizon

The other condition met by the Schwarzschild event horizon is $g^{rr} = 0$. For the Kerr metric, one has $g^{rr} = \Delta/\Sigma$, so that

$$g^{rr}(r,\theta) = 0 \iff \Delta(r) = r^2 - 2r_{\rm g}r + a^2 r_{\rm g}^2 = 0 \iff r = r_{\rm g}\left(1 \pm \sqrt{1-a^2}\right). \tag{4.13}$$

We are only interested in the positive root $r_{\rm h} = r_{\rm g}\left(1 + \sqrt{1-a^2}\right) \leq r_{\rm e}(\theta)$, which intersects the ergosphere at the poles $\theta = 0, \pi$ only (see Fig. 4.1). The hypersurface $\mathcal{H}$ defined by $r = r_{\rm h}$ is therefore a null surface. What is more, one can prove the following equality in Boyer-Lindquist coordinates: $g_{0\varphi}^2 - g_{00}g_{\varphi\varphi} = \Delta \sin^2\theta$. Consequently, within $r = r_{\rm h}$, Eq. (4.10) no longer has any real solution, and there can be no stationary observer. $\mathcal{H}$ is the true event horizon of a Kerr black hole (see the black circle in Fig. 4.1).

At the event horizon, we have $\Omega_- = \Omega_+ = -cg_{0\varphi}/g_{\varphi\varphi}$. As a result, the quantity

$$\omega_{\rm h} = -c\frac{g_{0\varphi}(r_{\rm h})}{g_{\varphi\varphi}(r_{\rm h})} = \frac{ac}{2r_{\rm h}} \tag{4.14}$$

is identified with the angular velocity of the black hole. For a low spin parameter $a \ll 1$ we have $\omega_{\rm h} \simeq (a/4)(c/r_{\rm g})$, whereas $\omega_{\rm h} \to c/(2r_{\rm g})$ for $a \to 1$.

An important takeway from this calculation is that the event horizon only exists if $|a| \leq 1$. A Kerr black hole with $|a| > 1$ would present a naked singularity, visible to distant observers. It is thought that a naked singularity cannot result from a regular initial state of the system through physically reasonable processes. This statement was formulated by R. Penrose as the weak cosmic censorship conjecture (see a discussion in Penrose 2002), and has not been proved so far. Therefore, we will consider $a = 1$ as the maximum possible value of the spin parameter.

4.1.5 Light surfaces

Kerr black holes exhibit another set of characteristic surfaces. Consider a point orbiting a Kerr black hole with angular velocity $\Omega \in [0, \omega_h]$: its rotation is timelike (subluminal) if the function f, defined in Eq. (4.10), is negative, and spacelike (superluminal) when f is positive. For a given Ω, the surfaces separating subluminal and superluminal rotation are the light surfaces, defined by $f(\Omega, r, \theta) = 0$. These light surfaces, parametrized by $r_l(\theta)$, will also have a special significance in this thesis, so it is useful to outline some of their properties. Their proofs can be found in Komissarov (2004a).

- There exist two light surfaces: the inner one is located between the ergosphere and the event horizon, whereas the outer one is outside the ergosphere.
- The inner and outer light surfaces never intersect.
- For the inner light surface, $r_l(\theta)$ increases as θ goes from 0 to $\pi/2$. The opposite holds for the outer light surface.

These surfaces are represented in Fig. 4.3. The outer light surface is very much similar to the light cylinder in pulsar magnetosphere theory, where Ω would be the pulsar's angular velocity. This can be seen in the non-relativistic limit: if $\Omega \ll c/r_g$ and $r \gg r_g$, the function f reduces to $f(\Omega, r, \theta) = -1 + \Omega^2 r^2 \sin^2\theta/c^2$. In this limit, the light surface has the equation

$$r_l(\theta) \sin\theta = \frac{c}{\Omega} : \tag{4.15}$$

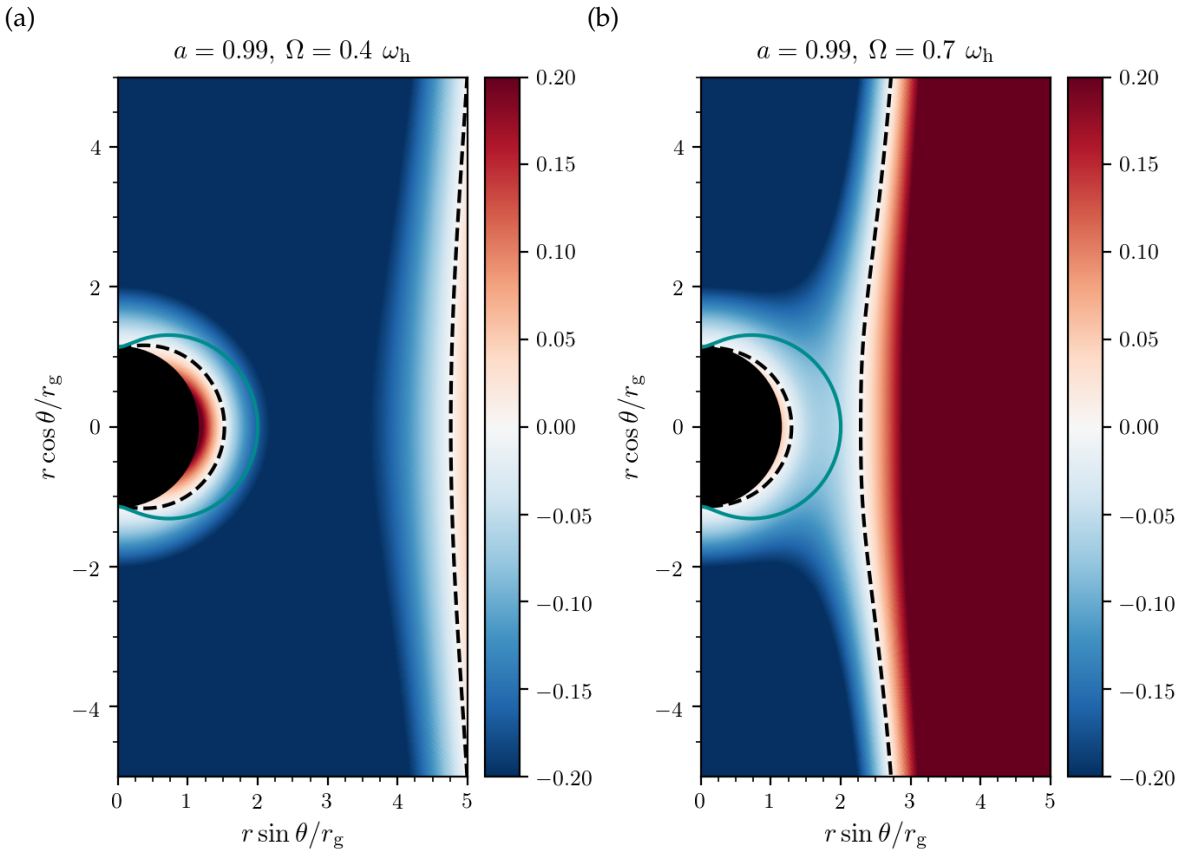


FIGURE 4.3: Spatial map of the light surface function f (Eq. (4.10)) for $a = 0.99$ and fixed $\Omega = 0.4\,\omega_h$ (a) and $\Omega = 0.7\,\omega_h$ (b).

this is precisely the equation of the cylinder of radius c/Ω. Outside the outer light surface, the rotation of the particle would be superluminal. Contrarily, the inner light surface is a peculiarity of Kerr black-hole magnetospheres, and has no classical counterpart. Inside the inner light surface, the motion of the particle would also be timelike (because $\Omega < \Omega_-$): the rotation of the particle is made somehow "too slow" because of the black hole's frame dragging. The importance of the light surfaces, and their relation to the Alfvén critical surfaces, will be described in Sec. 7.2.4.

4.2 Energy extraction

The goal of this section is to highlight another stark difference between Schwarzschild and Kerr black holes. To do so, we must first do a quick theoretical aside to define the conserved quantities of a test particle.

4.2.1 Killing vectors and conserved quantities

One of the most fecund ideas in theoretical physics is symmetries. In general relativity, this idea can be encompassed in the notion of Killing vectors, which generate isometries of the metric. A vector field K_μ is a Killing vector if it satisfies the condition

$$\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0. \tag{4.16}$$

In particular, if the components of the metric are all independent of a particular coordinate x, then $\boldsymbol{K} = \partial_x$ is a Killing vector associated with translational symmetry in the x direction. In this case, the system of coordinates is said to be adapted to the symmetries of the metric. This is the most straightforward way to derive Killing vectors, although in general it is impossible to find a coordinate system in which all Killing vectors have this form.

In classical physics, we know by Noether's theorem that symmetries of a system spawn the existence of conserved quantities. In this context, this is also the case: symmetries of the metric give rise to conserved charges, which can be defined thanks to the Killing vectors. Let K^μ be a given Killing vector field and $x^\mu(\tau)$ a geodesic, parametrized by its proper time τ, and let us define the charge $\mathcal{Q} = K_\mu \dot{x}^\mu$. Then $\mathcal{Q}$ is conserved along the geodesic.

Proof

It suffices to show that the directional derivative along the geodesic vanishes:

$$\begin{aligned}
\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\tau} &= \dot{x}^\mu \nabla_\mu \left(K_\nu \dot{x}^\nu\right) && (4.17)\\
&= \dot{x}^\mu K_\nu \nabla_\mu \dot{x}^\nu + \dot{x}^\mu \dot{x}^\nu \nabla_\mu K_\nu && (4.18)\\
&= \dot{x}^\mu \dot{x}^\nu \nabla_\mu K_\nu && (4.19)\\
&= \frac{1}{2} \dot{x}^\mu \dot{x}^\nu \left(\nabla_\mu K_\nu + \nabla_\nu K_\mu\right) = 0 && (4.20)
\end{aligned}$$

The third equality results from the geodesic equation $\dot{x}^\mu \nabla_\mu \dot{x}^\nu = 0$, and the final equality from the definition (4.16) of a Killing vector field.

Eq. (4.16) is equivalently linked to generating isometries of the metric and producing conserved charges. This can be understood intuitively. Loosely speaking, the metric is invariant in the direction given by K^μ, so a particle flowing in this direction experiences no "force" and conserves its momentum $\dot{x}^\mu$ along this direction.

The flat Minkowski spacetime has 10 independent Killing vectors, corresponding to a maximally symmetric space. The Kerr metric only possesses two straightforward Killing vectors $\boldsymbol{\eta} = \partial_t$ and $\boldsymbol{\psi} = \partial_\varphi$, since all its components are independent of t and φ. Therefore, for a given test particle (massive or massless) of momentum p^μ, one can define a conserved energy , associated with time translational invariance:

$$E = -cp^\mu \eta_\mu, \tag{4.21}$$

and a conserved angular momentum, associated with rotational invariance with respect to the black hole's spin axis:

$$L = p^\mu \psi_\mu. \tag{4.22}$$

If η^μ were the 4-velocity of a given observer, Eq. (4.21) would match the usual definition of the energy of a particle in special relativity. Thus, the conserved energy E is commonly denoted "energy-at-infinity", since it only achieves a physical meaning when measured by an observer at infinity with 4-velocity $\boldsymbol{\eta}$. Note that in general, E is not the energy that would be measured by a physical observer. In particular, inside the ergosphere, $\boldsymbol{\eta}$ is spacelike so it cannot be the 4-velocity of an observer. E is also sometimes called the "redshifted" energy.

4.2.2 Penrose process

As we saw, a Kerr black hole has the special property that the Killing vector field η^μ becomes spacelike within the ergosphere, although particles in the ergoregion can still make it to infinity. Outside the ergosphere, p^μ being timelike and future-oriented, E has to be positive, but this is no longer the case in the ergoregion: there can be particles of negative energy. Such thing would have no meaning for an observer at infinity, so particles of negative energy must unavoidably fall into the black hole ordrop their negative energy.

The existence of these particles in a region outside the event horizon makes the following process possible: a particle with momentum p_0^μ and energy-at-infinity E_0 could enter the ergosphere, then decay into a pair of particles with momenta p_1^μ and p_2^μ, so that particle 2 has negative energy-at-infinity $E_2 < 0$ (see Fig. 4.4). If particle 1 escapes from the ergosphere, its energy-at-infinity will be $E_1 = E_0 - E_2 > E_0$. Therefore, the outgoing particle has more energy than the ingoing one. The difference in energy must necessarily be provided by the black hole itself. Let us define the 4-vector $l^\mu = \eta^\mu + \Omega_\mathrm{h}\psi^\mu$; $\boldsymbol{l}$ is a null vector on the event horizon of the black hole. The scalar product of a future-oriented null vector with a future-oriented timelike vector is always negative (Gourgoulhon, 2010). Thus, we have $l^\mu p_{2\mu} = -E_2 + \Omega L_2 \leq 0 \Rightarrow L_2 \leq E_2/\Omega$. This implies that

$$L_1 - L_0 = -L_2 \geq -\frac{E_2}{\Omega} > 0. \tag{4.23}$$

Consequently, there has also been an extraction of angular momentum. Then, it is sensible to interpret this process as the extraction of rotational energy from the Kerr black hole. This provides a physical interpretation as to why no such process can be carried out for Schwarzschild black holes.

> From Eq. (4.23), one can deduce that the change in the black-hole mass M and angular momentum J verifies $\delta M \geq (\Omega/c^2)\delta J$. This is reminiscent of the second law of thermodynamics, which indicates a preferred direction of evolution for the system.

Unlike Schwarzschild black holes, it is possible to extract energy from a spinning black hole. Only a fraction of its mass is there for the taking, "reducible". The specific process that was previously outlined was first suggested by Penrose (Penrose & Floyd, 1971), but

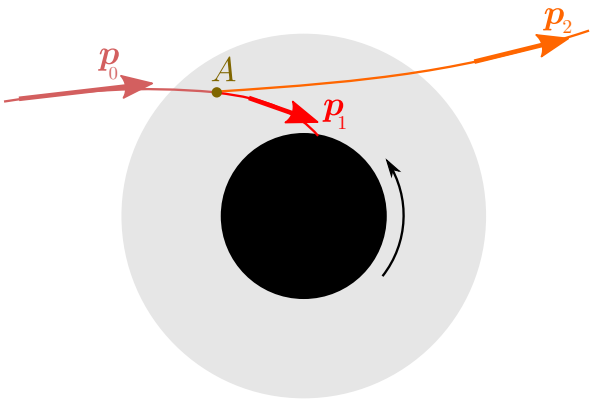


FIGURE 4.4: Illustration of the Penrose process in the $\theta = \pi/2$ equatorial plane. A particle with momentum $\boldsymbol{p}_0$ breaks up at A into two particles with momenta $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, with only one of which escaping to infinity. The gray shaded area represents the ergoregion. Figure adapted from Gourgoulhon (2021).

practical astrophysical implementations seem hard to devise. There is an analog of the Penrose process, known as "superradiance", which is based on wave amplification (Misner, 1972; Press & Teukolsky, 1972), but which also seems somehow implausible. The Blandford-Znajek mechanism is a more feasible way to extract energy from the black hole, by the means of electromagnetic fields.

4.A Appendix: Kerr-Schild coordinates

It has been underlined that the metric written in Boyer-Lindquist (hereafter, BL) coordinates presents a singularity at the event horizon. Just like in the Schwarzschild solution, this is merely a coordinate singularity. Nevertheless, such a singularity is inconvenient, for instance in the context of numerical simulations. It is possible to remove this coordinate singularity by a clever choice of coordinate transformation. The new coordinates, known as the Kerr-Schild (hereafter, KS) coordinates and denoted $(\tilde{t}, \tilde{r}, \tilde{\theta}, \tilde{\varphi})$, are the ones used in the general relativistic code `Zeltron` that I have used in my simulations (see Sec. 10). Recalling $\Delta = r^2 + 2rr_\mathrm{g} + a^2 r_\mathrm{g}^2$, KS coordinates are defined by the following transformation (Font et al., 1999; McKinney & Gammie, 2004):

$$\begin{cases} c\,\mathrm{d}\tilde{t} = c\,\mathrm{d}t + \dfrac{2r_\mathrm{g} r}{\Delta}\,\mathrm{d}r; & \text{(4.24a)} \\ \mathrm{d}\tilde{\varphi} = \mathrm{d}\varphi + \dfrac{a r_\mathrm{g}}{\Delta}\,\mathrm{d}r; & \text{(4.24b)} \\ \mathrm{d}\tilde{r} = \mathrm{d}r; & \text{(4.24c)} \\ \mathrm{d}\tilde{\theta} = \mathrm{d}\theta. & \text{(4.24d)} \end{cases}$$

Integrating Eq. (4.24a) and (4.24b) provides the full expression of the new KS coordinates:

$$\begin{cases} c\tilde{t} = ct + r_\mathrm{g} \ln(\Delta) + \dfrac{r_\mathrm{g}}{\sqrt{1-a^2}} \ln\left(\dfrac{r - r_\mathrm{g}\left(1+\sqrt{1-a^2}\right)}{r - r_\mathrm{g}\left(1-\sqrt{1-a^2}\right)} \right), & \text{(4.25a)} \\ \tilde{\varphi} = \varphi + \dfrac{a}{2\sqrt{1-a^2}} \ln\left(\dfrac{r - r_\mathrm{g}\left(1+\sqrt{1-a^2}\right)}{r - r_\mathrm{g}\left(1-\sqrt{1-a^2}\right)} \right), & \text{(4.25b)} \end{cases}$$

along with $\tilde{r} = r$ and $\tilde{\theta} = \theta$. The Jacobian matrix $J^{\tilde{\mu}}{}_{\nu} = \left(\partial \tilde{x}^{\tilde{\mu}} / \partial y^{\nu}\right)$ and its inverse are given by

$$J^{\tilde{\mu}}{}_{\nu} = \begin{pmatrix} 1 & 2r_{\mathrm{g}} r/\Delta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & a r_{\mathrm{g}}/\Delta & 0 & 1 \end{pmatrix}; \quad \left(J^{-1}\right)^{\mu}{}_{\tilde{\nu}} = \begin{pmatrix} 1 & -2r_{\mathrm{g}} r/\Delta & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -a r_{\mathrm{g}}/\Delta & 0 & 1 \end{pmatrix}. \tag{4.26}$$

In these KS coordinates, the metric reads (Komissarov, 2004a; Takahashi, 2007):

$$\mathrm{d}s^2 = g_{\tilde{0}\tilde{0}} c^2 \,\mathrm{d}\tilde{t}^2 + 2 g_{\tilde{0}\tilde{\varphi}} c \,\mathrm{d}\tilde{t}\,\mathrm{d}\tilde{\varphi} + 2 g_{\tilde{0}\tilde{r}} c \,\mathrm{d}\tilde{t}\,\mathrm{d}\tilde{r} + g_{\tilde{r}\tilde{r}} \,\mathrm{d}\tilde{r}^2 + g_{\tilde{\theta}\tilde{\theta}} \,\mathrm{d}\tilde{\theta}^2 + g_{\tilde{\varphi}\tilde{\varphi}} \,\mathrm{d}\tilde{\varphi}^2 + 2 g_{\tilde{r}\tilde{\varphi}} \,\mathrm{d}\tilde{r}\,\mathrm{d}\tilde{\varphi}, \tag{4.27}$$

where

$$g_{\tilde{0}\tilde{0}} = g_{00} = z - 1; \quad g_{\tilde{0}\tilde{\varphi}} = g_{0\varphi} = -a r_{\mathrm{g}} z \sin^2\theta; \quad g_{\tilde{0}\tilde{r}} = z; \tag{4.28}$$

$$g_{\tilde{r}\tilde{r}} = 1 + z; \quad g_{\tilde{\theta}\tilde{\theta}} = g_{\theta\theta} = \Sigma; \quad g_{\tilde{\varphi}\tilde{\varphi}} = g_{\varphi\varphi} = \frac{A \sin^2\theta}{\Sigma}; \quad g_{\tilde{r}\tilde{\varphi}} = -a r_{\mathrm{g}} \sin^2\theta \,(1 + z)\,. \tag{4.29}$$

For conciseness, we have defined $z = 2 r_{\mathrm{g}} r / \Sigma$. The inverse components of the metric in the KS coordinates are (Takahashi, 2007)

$$g^{\tilde{0}\tilde{0}} = 1 + z; \quad g^{\tilde{0}\tilde{\varphi}} = 0; \quad g^{\tilde{0}\tilde{r}} = z \tag{4.30}$$

$$g^{\tilde{r}\tilde{r}} = \frac{\Delta}{\Sigma}; \quad g^{\tilde{\theta}\tilde{\theta}} = \frac{1}{\Sigma}; \quad g^{\tilde{\varphi}\tilde{\varphi}} = \frac{1}{\Sigma \sin^2\theta}; \quad g^{\tilde{r}\tilde{\varphi}} = \frac{a r_{\mathrm{g}}}{\Sigma}. \tag{4.31}$$

A useful identity is $h^{r\varphi} h_{\varphi\varphi} + h_{r\varphi} h^{rr} = 0$. This transformation has allowed us to get rid of the coordinate singularity at $r = r_{\mathrm{h}}$ (which was due to the presence of Δ), both in $g_{\tilde{\mu}\tilde{\nu}}$ and in $g^{\tilde{\mu}\tilde{\nu}}$. Unfortunately, the coordinates $(\tilde{r}, \tilde{\theta}, \tilde{\varphi})$ are no longer orthogonal, so that extra care is needed in order to interpret quantities in the KS coordinates. One can check that the determinant $\tilde{g}$ of the metric in these coordinates has the same expression than in BL coordinates:

$$\sqrt{-\tilde{g}} = \Sigma \sin\theta = \sqrt{-g}. \tag{4.32}$$

Under this change of coordinates, a 4-vector X^{μ} is transformed as follows (using $X^{\tilde{\mu}} = J^{\tilde{\mu}}{}_{\nu} X^{\nu}$):

$$\begin{cases} X^{\tilde{0}} = X^0 + \dfrac{2 r_{\mathrm{g}} r}{\Delta} X^r & \text{(4.33a)} \\ X^{\tilde{r}} = X^r & \text{(4.33b)} \\ X^{\tilde{\theta}} = X^{\theta} & \text{(4.33c)} \\ X^{\tilde{\varphi}} = X^{\varphi} + \dfrac{a r_{\mathrm{g}}}{\Delta} X^r & \text{(4.33d)} \end{cases}$$

Likewise, a 1-form X_{μ} transforms as follows (using $X_{\tilde{\mu}} = \left(J^{-1}\right)^{\mu}{}_{\tilde{\nu}} X_{\nu}$):

$$\begin{cases} X_{\tilde{0}} = X_0 & \text{(4.34a)} \\ X_{\tilde{r}} = X_r - \dfrac{2 r_{\mathrm{g}} r}{\Delta} X_0 - \dfrac{a r_{\mathrm{g}}}{\Delta} X_{\varphi} & \text{(4.34b)} \\ X_{\tilde{\theta}} = X_{\theta} & \text{(4.34c)} \\ X_{\tilde{\varphi}} = X_{\varphi} & \text{(4.34d)} \end{cases}$$

Only the r component of a 1-form is transformed. Similarly, a $(0,2)$ tensor $F_{\mu\nu}$ transforms as $F_{\tilde{\mu}\tilde{\nu}} = \left(J^{-1}\right)^{\mu}_{\ \tilde{\mu}} \left(J^{-1}\right)^{\nu}_{\ \tilde{\nu}} F_{\mu\nu}$. Inserting Eq. (4.26), it follows that all components $F_{\mu\nu}$ for $\mu \neq r$ and $\nu \neq r$ are invariant under the coordinate transformation. The light surface function f, defined in Eq. (4.10), has the same expression in BL and KS coordinates, because its definition does not involve any component of the metric with r.

Chapter 5

$3+1$ electrodynamics

The goal of this chapter is to derive a convenient formulation of electrodynamics in curved spacetime. Carrying out this task requires a different formulation of general relativity. The classical formulation is usually expressed in a frame-independent way, and deals with 4-dimensional objects to represent physical quantities. This is very much the spirit of the theory, and it is highly justified when dealing with dynamical spacetimes. Unfortunately, its connection with an already complex theory makes it harder to rely on physical intuition. By contrast, the $3+1$ formulation of electrodynamics, introduced by Thorne & MacDonald (1982) and MacDonald & Thorne (1982), introduces an "absolute" 3-dimensional space and an "absolute" time. This makes a connection with classical physics much easier to establish. Not only does this decomposition make electrodynamics more tractable, with 3-vectors representing the electromagnetic fields, but it is also very well suited for computational purposes, as quantities can be evolved with respect to this absolute time. We will see in Chap. 9 and 10 how this formalism allows us to rely on the architecture of codes initially designed in flat spacetime. This approach is appropriate when one deals with stationary spacetime, as is the case in this thesis.

> The $3+1$ formalism of general relativity was first developed by mathematicians, such as Yvonne Choquet-Bruhat in the 1950's. Nowadays, this is the main approach in solving Einstein's equations numerically.

5.1 $3+1$ formulation of general relativity

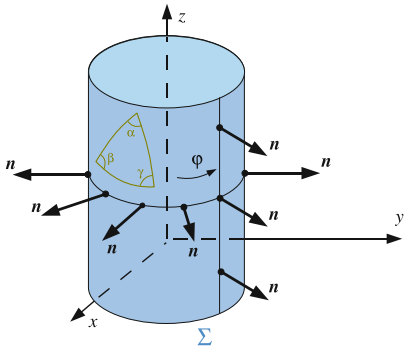
Most of the material in this section is covered in a very thorough way in the exhaustive textbook by Gourgoulhon (2012). A simpler introduction can be found in Dodin & Fisch (2010). In this section, Greek indices span from 0 to 3 and Latin indices from 1 to 3. Vectors will be written in bold, unlike 1-forms. In the following, we assume a stationary spacetime.

5.1.1 Hypersurfaces

First, we focus on how to switch from 4-dimensional objects, defined in a 4-manifold, to 3-dimensional objects, defined on spacelike hypersurfaces. These spacelike hypersurfaces can constitute what can be thought of as "space". Let us define a 3-dimensional spacelike hypersurface Σ. The proper definition of a hypersurface is related to the involved notion of "embedding": that is, how a given 3-manifold is defined with respect to the 4-dimensional spacetime. In particular, a hypersurface can be intrinsically flat (vanishing Riemann tensor), but have an extrinsic curvature due to the way it is embedded in spacetime. In the following, we will pass over these mathematical sophistications.

Example

Let us simply provide an example to give the reader a sense of this notion: a cylinder of radius R and symmetry axis (Oz) in $\mathbb{R}^3$ is defined by the equation $\sqrt{x^2+y^2} = R$. The cylindrical coordinates φ and z map entirely the cylinder. In this system, the line element on the cylinder reads $\mathrm{d}s^2 = R^2\,\mathrm{d}\varphi^2 + \mathrm{d}z^2$, so that the metric is flat. Yet, the cylinder has extrinsic curvature due to how its normal vector varies along its surface.



We assume that Σ is the level surface of some scalar field t, then by definition the 1-form ∇t is normal to Σ. Σ is spacelike if the dual vector $\boldsymbol{\nabla} t$ is timelike. In that case, let us define $\boldsymbol{n} = -\alpha \boldsymbol{\nabla} t$ as a unit normal vector of Σ, such that $n^\mu n_\mu = -1$. All 4-vectors can be decomposed as the sum of a vector in Σ and one colinear to $\boldsymbol{n}$. Let us introduce the symmetric tensor

$$h^\mu{}_\nu = \delta^\mu{}_\nu + n^\mu n_\nu, \tag{5.1}$$

where $\delta^\mu{}_\nu$ is the Kronecker tensor. h is the orthogonal projector on Σ. Indeed, it satisfies $h^\mu{}_\nu n^\nu = 0$, and $h^\mu{}_\nu X^\nu = X^\mu$ for all vector X tangent to Σ.

Let us also introduce three arbitrary vectors $\boldsymbol{e}_i$ forming an orthonormal basis of Σ, so that $(\boldsymbol{n}, \boldsymbol{e}_i)$ forms an orthonormal basis of the spacetime. We define the projection $\overline{X}^\mu = h^\mu{}_\nu \cdot X^\nu$ of a 4-vector $\boldsymbol{X}$: its component along $\boldsymbol{n}$ is zero. Because it is a tensor contraction, $\overline{X}^\mu$ must transform as vector components, and so do the $\overline{X}^i$ components. Therefore, $\overline{\boldsymbol{X}} = \overline{X}^i \boldsymbol{e}_i$ can be considered as a spatial 3-vector in Σ.

Likewise, this projector can be extended to act upon 1-forms and tensors. In particular, the line element $\mathrm{d}\overline{\boldsymbol{x}} \cdot \mathrm{d}\overline{\boldsymbol{x}}$ in Σ reads

$$\begin{aligned}
\mathrm{d}\overline{\boldsymbol{x}} \cdot \mathrm{d}\overline{\boldsymbol{x}} &= \mathrm{d}\overline{x}^i\,\mathrm{d}\overline{x}_i && (5.2)\\
&= \mathrm{d}\overline{x}^\mu\,\mathrm{d}\overline{x}_\mu && (5.3)\\
&= h_{\mu\nu}\,\mathrm{d}\overline{x}^\mu\,\mathrm{d}\overline{x}^\nu && (5.4)\\
&= h_{ij}\,\mathrm{d}\overline{x}^i\,\mathrm{d}\overline{x}^j, && (5.5)
\end{aligned}$$

with $h_{\mu\nu} = g_{\mu\nu} h^\mu{}_\nu$. The first equality can be seen as the definition of the line element, for a small 3-vector element $\mathrm{d}x$ in Σ; the second equality is due to the component along $\boldsymbol{n}$ of $\mathrm{d}\overline{\boldsymbol{x}}$ being zero; the third equality consisted in lowering indices for the h operator. Therefore, the spatial part of the orthogonal projector h can be seen as the spatial part of the metric on the hypersurface Σ. Formally, it is the metric "induced" on Σ by the 4-metric g, where the covariant components of h are

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu. \tag{5.6}$$

Now the hypersurface is endowed with a metric, induced by the properties of the global spacetime.

5.1.2 Foliation of spacetime

In the last section we have focused on the geometry of a single hypersurface. The key idea of the $3+1$ formalism is to define a foliation of spacetime by a continuous set of hypersurfaces. We define a regular scalar field t on the spacetime (such that $\boldsymbol{\nabla} t \cdot \boldsymbol{\nabla} t \neq 0$ everywhere), and define each hypersurface Σ_t as a spacelike level surface of t. Different hypersurfaces of the foliation cannot intersect because t is regular (see Fig. 5.1).

Now we need to introduce coordinates on the spacetime. First we define a coordinate system (x^i) which varies smoothly between different hypersurfaces. The associated coordinate basis is $\boldsymbol{e}_i = \partial_i$, where the $\boldsymbol{e}_i$ are tangent to the hypersurfaces. In addition to these, we add a fourth coordinate $x^0 = t$ associated with the time vector $\boldsymbol{e}_0 = \partial_t$. We have already defined the unit vector $\boldsymbol{n} = -\alpha \boldsymbol{\nabla} t$, which is timelike, future-oriented and normal to the hypersurface. α is generally called the "lapse" function, for reasons which will appear later. In order to complete the description of the foliation, one has to parametrize this vector $\boldsymbol{e}_0$. Importantly, $\boldsymbol{e}_0$ is not necessarily timelike. By construction, it is only required to remain tangent to lines of constant x^i.

Any infinitesimal 4-vector can be written as $\mathrm{d}\boldsymbol{x} = \mathrm{d}x^i \boldsymbol{e}_i + \mathrm{d}t \boldsymbol{e}_0$. Because $\boldsymbol{\nabla} t$ is normal to Σ_t, this yields $\boldsymbol{\nabla} t \cdot \mathrm{d}\boldsymbol{x} = (\boldsymbol{\nabla} t \cdot \boldsymbol{e}_0)\, \mathrm{d}t$. But since we also have $\mathrm{d}t = \boldsymbol{\nabla} t \cdot \mathrm{d}\boldsymbol{x}$, both equations combine to yield $\boldsymbol{\nabla} t \cdot \boldsymbol{e}_0 = 1$, which amounts to saying that the 1-form $\boldsymbol{\nabla} t$ is dual to the vector ∂_t. Given a choice of lapse function α, we deduce that

$$\boldsymbol{n} \cdot \boldsymbol{e}_0 = -\alpha, \tag{5.7}$$

which is the only requirement on $\boldsymbol{e}_0$. The time vector $\boldsymbol{e}_0 = \partial_t$ does not have to be normal to Σ, as long as Eq.(5.7) is satisfied. All in all, the general form of the time vector is

$$\boldsymbol{e}_0 = \alpha \boldsymbol{n} + \boldsymbol{\beta}, \tag{5.8}$$

where $\boldsymbol{\beta}$ is tangent to Σ_t (because from Eq. (5.7) we have $\boldsymbol{n} \cdot \boldsymbol{\beta} = 0$). $\boldsymbol{\beta}$ was coined the "shift" vector. From Eq. (5.8), we see that $\boldsymbol{\beta} = \beta^i \boldsymbol{e}_i$ is the projection of $\boldsymbol{e}_0$ onto Σ. Also, the square norm of $\boldsymbol{e}_0$ is $\boldsymbol{e}_0 \cdot \boldsymbol{e}_0 = -\alpha^2 + \boldsymbol{\beta}^2$, so that $\boldsymbol{e}_0$ does not have to be timelike (as pointed out earlier). The lapse function α and the shift vectors $\boldsymbol{\beta}$ are used to parametrize the foliation of spacetime, and thus specify the metric. The geometry of the foliation is represented in Fig. 5.2.

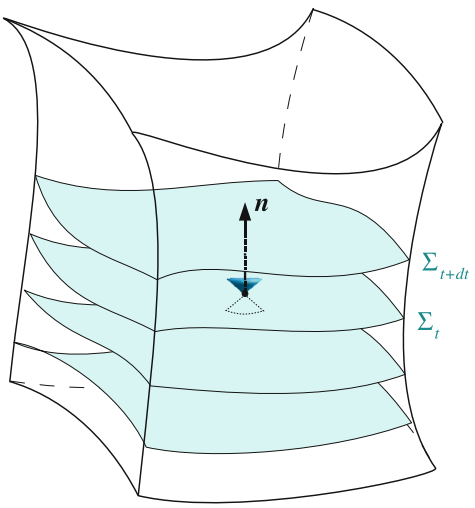


FIGURE 5.1: Sketch of the foliation of spacetime into hypersurfaces Σ_t of constant t. $\boldsymbol{n}$ is a unit vector normal to the hypersurface Σ_t. Figure adapted from Gourgoulhon (2012).

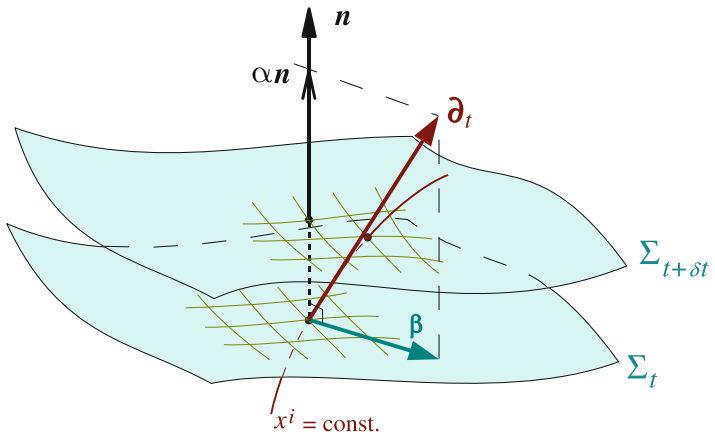


FIGURE 5.2: Geometry of the foliation. The red solid line is a line of constant coordinates x^i. The time vector $\boldsymbol{e}_0 = \partial_t$ is tangent to this line. Figure adapted from Gourgoulhon (2012).

5.1.3 Interpretation of the lapse function and shift vectors

$\boldsymbol{n}$ is a timelike vector, so it can be considered as the normalized 4-velocity of a local observer. Such observers are denoted "fiducial observers" (FIDOs). The $3+1$ formalism has the benefit of naturally introducing a class of privileged observers. Since $\boldsymbol{n}$ is normal to Σ_t, this hypersurface is locally the set of simultaneous events of this observer.

Eq. (5.8) allows to write down the components of $\boldsymbol{n}$ in the coordinate basis $(\boldsymbol{e}_\alpha) = (\partial_\alpha)$:

$$n^\mu = \left(\frac{1}{\alpha}, -\frac{\beta^i}{\alpha}\right). \tag{5.9}$$

The covariant components of $\boldsymbol{n}$ (or equivalently, the components of its 4-momentum) in the dual basis $\mathrm{d}x^\alpha$ are obtained similarly, by noting that $\boldsymbol{n} = -\alpha\boldsymbol{\nabla}t \Rightarrow n = -\alpha\nabla t$, so that

$$n_\mu = (-\alpha, 0_i). \tag{5.10}$$

Let us consider two points M and M' on the worldline of a given FIDO, on two hypersurfaces labeled by t and $t' = t + \delta t$ respectively. To evaluate the proper time $\delta\tau$ measured by the FIDO between these two events, we need to find the length of the infinitesimal vector $\delta\boldsymbol{x}$ separating M and M'. This vector must be colinear to the 4-velocity of the FIDO, so we write $\delta\boldsymbol{x} = N\delta t\boldsymbol{n}$, with N to be determined. The difference in t can be rewritten as

$$t' - t = \delta t = \boldsymbol{\nabla}t \cdot \delta\boldsymbol{x} = N\boldsymbol{\nabla}t \cdot \boldsymbol{n}\delta t. \tag{5.11}$$

But we know that $\boldsymbol{n} = -\alpha\boldsymbol{\nabla}t$ and $\alpha^2 = -\boldsymbol{\nabla}t \cdot \boldsymbol{\nabla}t$, so that this equation yields $N = \alpha$. Therefore, the infinitesimal proper time measured by the FIDO is

$$\delta\tau_{\mathrm{FIDO}} = \sqrt{-g_{\mu\nu}\delta x^\mu x^\nu}/c^2 = \alpha\delta t. \tag{5.12}$$

This justifies the name "lapse function" given to α, relating the physical proper time with the coordinate time t. In this formalism, t can be seen as a universal "absolute" time, parametrizing the hypersurfaces Σ_t, which define an "absolute" space.

Likewise, one can see from Eq. (5.9) that the shift vector $\boldsymbol{\beta}$ can be regarded as the FIDO's spatial velocity (up to a proportionality factor). We also note that the spatial components of the FIDO's momentum are zero: the FIDO can be seen as at rest with respect to the absolute space. Then, β^i is rather the rate at which the coordinate system on each hypersurface (the lines of constant x^i) is shifting with respect to the FIDO's wordline.

In particular, since we have $\boldsymbol{n}\cdot\boldsymbol{\psi}=0$, FIDOs have vanishing angular momentum: they are also "zero-angular momentum observers" (ZAMOs). Nevertheless, FIDOs are not inertial observers, their 4-acceleration being non-zero.

5.1.4 Metric

Finally, since α and $\boldsymbol{\beta}$ fully determine the metric, it is useful to write the components of the metric with this choice of foliation. Each component can be computed as follows:

$$g_{00}=g(\boldsymbol{e}_0,\boldsymbol{e}_0)=\boldsymbol{e}_0\cdot\boldsymbol{e}_0=-\alpha^2+\boldsymbol{\beta}\cdot\boldsymbol{\beta} \tag{5.13}$$
$$g_{0i}=g(\boldsymbol{e}_0,\boldsymbol{e}_i)=(\alpha\boldsymbol{n}+\boldsymbol{\beta})\cdot\boldsymbol{e}_i=\boldsymbol{\beta}\cdot\boldsymbol{e}_i=\beta_i \tag{5.14}$$
$$g_{ij}=g(\boldsymbol{e}_i,\boldsymbol{e}_j)=h_{ij}. \tag{5.15}$$

For compactness, we can write

$$g_{\mu\nu}=\begin{pmatrix}-\alpha^2+\boldsymbol{\beta}^2 & \beta_j\\ \beta_i & h_{ij}\end{pmatrix}. \tag{5.16}$$

The line element can be written as

$$\mathrm{d}s^2=\left(\boldsymbol{\beta}^2-\alpha^2\right)c^2\,\mathrm{d}t^2+2\beta_i\,\mathrm{d}x^i c\,\mathrm{d}t+h_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j=-\alpha^2c^2\,\mathrm{d}t^2+h_{ij}\left(\mathrm{d}x^i+\beta^i c\,\mathrm{d}t\right)\left(\mathrm{d}x^j+\beta^j c\,\mathrm{d}t\right), \tag{5.17}$$

which is convenient to determine quickly the lapse function and shift vector if the metric is given. The inverse metric components are:

$$g^{\mu\nu}=\begin{pmatrix}-1/\alpha^2 & \beta^j/\alpha^2\\ \beta^i/\alpha^2 & h^{ij}-\beta^i\beta^j/\alpha^2\end{pmatrix}, \tag{5.18}$$

where h^{ij} is the inverse matrix of h_{ij}.

We will also derive a useful relationship between the determinants of the 4-metric g and the spatial 3-metric h. Note that g and h depend on the choice of coordinates, and are not scalar quantities. By Cramer's rule, the $(0,0)$ element of the inverse metric is related to the minor $(0,0)$ M_{00} of $g_{\mu\nu}$ by $g^{00}=M_{00}/g$. But thanks to Eq. (5.16), M_{00} is also the determinant of the 3-metric h, so that $g^{00}=h/g$. Using the explicit value of g^{00}, we find

$$\sqrt{-g}=\alpha\sqrt{h}. \tag{5.19}$$

In BL coordinates, the shift vector is purely azimuthal: we get from $\beta_i=g_{0i}$ that

$$\boldsymbol{\beta}_{\mathrm{BL}}=\beta^\varphi_{\mathrm{BL}}\partial_\varphi,\quad \beta^\varphi_{\mathrm{BL}}=\frac{g_{0\varphi}}{g_{\varphi\varphi}}. \tag{5.20}$$

The BL FIDO has a purely azimuthal motion relative to the coordinate grid, with a constant angular velocity $\omega=-c\beta^\varphi$. From $g^{00}=-1/\alpha^2$, we obtain the lapse function:

$$\alpha_{\mathrm{BL}}=\sqrt{\frac{\Delta\Sigma}{A}}. \tag{5.21}$$

On the other hand, in KS coordinates, the lapse function takes a simpler form:

$$\alpha_{\mathrm{KS}}=\frac{1}{\sqrt{1+z}}, \tag{5.22}$$

where $z = 2r_g r/\Sigma$. Because $\beta^i = g^{0i}\alpha^2$, we see that the shift vector is purely radial:

$$\boldsymbol{\beta}_{\mathrm{KS}} = \beta^r_{\mathrm{KS}}\partial_r, \quad \beta^r_{\mathrm{BL}} = \frac{z}{1+z}. \tag{5.23}$$

The KS FIDO slides radially toward the true singularity at $r = 0$. Because this physical observer is able to cross the event horizon, there can be no coordinate singularity at the event horizon as seen by the FIDO (see the Appendix 4.A). Note however that $\beta_\varphi \neq 0$.

5.2 Maxwell's equations in general relativity

We can start by reviewing Maxwell's equations in special relativity, then promote them to a generally covariant form by minimal coupling.

> Basically, to obtain equations which are valid in any gravitational field, one must write down Lorentz invariant equations, then promote them by replacing the Minkowski metric by the spacetime metric and partial derivatives by covariant derivatives, in order to get fully covariant equations. This is known as the minimal coupling procedure.

The basic dynamical field is the vector potential A_μ, which allows us to define the Maxwell field strength (Blau, 2018)

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \nabla_\mu A_\nu - \nabla_\nu A_\mu, \tag{5.24}$$

where the first equality is the definition of $F_{\mu\nu}$ and the second equality results from the spacetime having no torsion. The generally covariant Maxwell's equations are (Gourgoulhon, 2010; Blau, 2018)

$$\begin{cases} \nabla_\mu F^{\nu\mu} = \dfrac{4\pi}{c} I^\nu & \text{(5.25a)} \\ \nabla_\mu {*F}^{\nu\mu} = 0, & \text{(5.25b)} \end{cases}$$

where ${*F}$ is the Hodge dual of the Maxwell tensor, whose components are defined by ${*F}_{\alpha\beta} = (1/2)e_{\alpha\beta\mu\nu}F^{\mu\nu}$ and ${*F}^{\alpha\beta} = (1/2)e^{\alpha\beta\mu\nu}F_{\mu\nu}$. We have used the Levi-Civita tensor $e_{\alpha\beta\mu\nu}$ introduced in the Appendix 5.A.1. I^ν is the 4-electric current vector. Eq. (5.25b) is also often written into the (less compact) form $\nabla_\alpha F_{\beta\gamma} + \nabla_\beta F_{\gamma\alpha} + \nabla_\gamma F_{\alpha\beta} = 0$, whereas Eq. (5.25a) implies charge conservation $\nabla_\mu I^\mu = 0$.

Proof

Indeed, we have

$$\nabla_\mu F^{\nu\mu} = \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}F^{\nu\mu}\right) \Rightarrow \nabla_\nu I^\nu = \frac{1}{4\pi\sqrt{-g}}\partial_\nu\partial_\mu\left(\sqrt{-g}F^{\nu\mu}\right) = 0, \tag{5.26}$$

where the last equality comes from the antisymmetry of $F^{\mu\nu}$. In this equation, we have used the identities (5.80) and (5.81).

The covariant equivalent of the Lorenz gauge for the vector potential is $\nabla_\mu A^\mu = 0$. Written in terms of A^μ, Maxwell's equations now read

$$\nabla_\mu F^{\mu\nu} = \nabla_\mu\left(\nabla^\mu A^\nu - \nabla^\nu A^\mu\right) = \Box A^\nu - R^\nu{}_\mu A^\mu = \frac{4\pi}{c} I^\nu, \tag{5.27}$$

where $R_{\mu\nu}$ is the Ricci tensor of the metric and $\Box = \nabla_\mu \nabla^\mu$ is the d'Alembert ("box") operator. This reformulation of Maxwell's equations will be used to find an analytical solution for the electromagnetic field around a Kerr black hole.

It is interesting to underline the asymmetry between the two Maxwell's equations (5.25a) and (5.25b): one involves a source term (electric charges) while the other one does not, due to the absence of magnetic charges in nature. However, they become symmetrical in the absence of source terms. The duality between the two equations is performed by the action of the Hodge operator $*$. As a result, the Hodge dual $*F_{\mu\nu}$ of a solution of the vacuum Maxwell's equations $F_{\mu\nu}$ is also a solution of these equations. Basically, going from $F_{\mu\nu}$ to $*F_{\mu\nu}$ amounts to switching the electric and magnetic fields. We will make use of this fact in Sec. 6.1.2 in order to derive the expression for a magnetic monopole in Kerr spacetime. Let us also recall that for a p-form ϖ, we have $**\varpi = (-1)^{p+1}\varpi$, so that $**F = -F$.

5.3 $3+1$ reformulation of Maxwell's equations

This reformulation of Maxwell's equations has been introduced in Komissarov (2004a).

5.3.1 Derivation

As promised in the introduction of this chapter, we will rewrite the generally covariant Maxwell's equations (5.25a) and (5.25b) in the $3+1$ formalism. The coordinates (x^μ) that we use to map the spacetime are now (t, x^i), with t a universal time coordinate and (x^i) a coordinate system adapted to the foliation. Using the identities (5.80) and (5.19), Eq. (5.25b) can be split into a time part and a spatial part:

$$\partial_i \left(\alpha \sqrt{h} * F^{ti} \right) = 0 \tag{5.28}$$

$$\frac{1}{c\sqrt{h}} \partial_t \left(\alpha \sqrt{h} * F^{it} \right) + \frac{1}{\sqrt{h}} \partial_j \left(\alpha \sqrt{h} * F^{ij} \right) = 0. \tag{5.29}$$

Let us define the spatial 3-vectors $B^i = \alpha * F^{i0}$ and $E^i = h^{ij} E_i$, with $E_i = (1/2)\alpha e_{ijk} * F^{jk}$. The covariant component E_i can actually be rewritten as $E_i = F_{i0}$.

Proof

$$E_i = \frac{1}{2} \alpha \sqrt{h} \varepsilon_{ijk} * F^{jk} \tag{5.30}$$

$$= \frac{1}{2} \alpha \sqrt{h} \varepsilon_{ijk} \frac{1}{2\sqrt{-g}} \varepsilon^{\mu\nu jk} F_{\mu\nu} \tag{5.31}$$

$$= \frac{1}{2} \varepsilon_{ijk} \left(\varepsilon^{0ljk} F_{0l} + \varepsilon^{l0jk} F_{l0} \right) \tag{5.32}$$

$$= \frac{1}{2} \varepsilon_{ijk} \varepsilon^{jkl} F_{l0} = \delta_i{}^l F_{l0} = F_{i0}, \tag{5.33}$$

where we have used the contraction (5.73).

Similarly, B^i could be rewritten as $B^i = (1/2) e^{ijk} F_{jk}$. Using the identity (5.80), Eq. (5.28) can be recast in the form $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$. The first term in Eq. (5.29) is the i component of the

vector $\partial_t\left(\sqrt{h}\boldsymbol{B}\right)/c\sqrt{h}$, whereas the second term can be rewritten as

$$\frac{1}{\sqrt{h}}\partial_j\left(\alpha\sqrt{h}{*F}^{ij}\right)=\frac{1}{2\sqrt{h}}\varepsilon^{ij\mu\nu}\partial_j F_{\mu\nu}=\frac{1}{2\sqrt{h}}\left(\varepsilon^{ij0k}\partial_j F_{0k}+\varepsilon^{ijk0}\partial_j F_{k0}\right)=\frac{1}{\sqrt{h}}\varepsilon^{ijk}\partial_j F_{k0}. \tag{5.34}$$

Thus, the second term is actually the i component of the curl $\boldsymbol{\nabla}\times\boldsymbol{E}$ (see the definition (5.71)), so that the spatial part of Eq. (5.25b) actually reads

$$\frac{1}{\sqrt{h}}\frac{\partial\sqrt{h}\boldsymbol{B}}{\partial t}+c\boldsymbol{\nabla}\times\boldsymbol{E}=\boldsymbol{0}. \tag{5.35}$$

The same procedure applies to Eq. (5.25a), so we will simply state the result. We define the 3-vectors $D^i=\alpha F^{0i}$ and $H^i=h^{ij}H_j$, with $H_i=(1/2)\alpha e_{ijk}F^{jk}$, and the charge density $\rho=\alpha I^0/c=\alpha I^t$ and current density $J^i=\alpha I^i$. The time part of Eq. (5.25a) gives $\boldsymbol{\nabla}\cdot\boldsymbol{D}=4\pi\rho$, whereas the spatial part is the i component of

$$\frac{1}{\sqrt{h}}\frac{\partial\sqrt{h}\boldsymbol{D}}{\partial t}=c\boldsymbol{\nabla}\times\boldsymbol{H}-4\pi\boldsymbol{J}. \tag{5.36}$$

If we further assume the metric is stationary, we end up with the following set of $3+1$ Maxwell's equations:

$$\begin{cases}\boldsymbol{\nabla}\cdot\boldsymbol{D}=4\pi\rho, & \text{(5.37a)}\\ \dfrac{\partial\boldsymbol{D}}{\partial t}=c\boldsymbol{\nabla}\times\boldsymbol{H}-4\pi\boldsymbol{J}, & \text{(5.37b)}\\ \boldsymbol{\nabla}\cdot\boldsymbol{B}=0, & \text{(5.37c)}\\ \dfrac{\partial\boldsymbol{B}}{\partial t}=-c\boldsymbol{\nabla}\times\boldsymbol{E}. & \text{(5.37d)}\end{cases}$$

The $3+1$ version of the charge conservation equation is obtained by applying the divergence operator to Eq. (5.37b), which yields using Eq. (5.37a) the usual law: $\partial_t\rho+\boldsymbol{\nabla}\cdot\boldsymbol{J}=0$. The 4-vector potential can also be split into a time and spatial part. Let us define the scalar potential $\phi=-A_0$ and the 3-vector potential $\mathcal{A}^i=h^{ij}A_j$. Then, using the definition of the Maxwell field strength, we have

$$E_i=F_{i0}=\partial_i A_0-\frac{1}{c}\partial_t A_i\Rightarrow\boldsymbol{E}=-\frac{1}{c}\partial_t\boldsymbol{\mathcal{A}}-\boldsymbol{\nabla}\phi, \tag{5.38}$$

$$B^i=\frac{1}{2}e^{ijk}F_{jk}=\frac{1}{2\sqrt{h}}\varepsilon^{ijk}\left(\partial_j A_k-\partial_k A_j\right)=\frac{1}{\sqrt{h}}\varepsilon^{ijk}\partial_j A_k\Rightarrow\boldsymbol{B}=\boldsymbol{\nabla}\times\boldsymbol{\mathcal{A}}. \tag{5.39}$$

It is instructive to show that just like in flat spacetime, if the electromagnetic fields are invariant by rotation around the z axis, A_φ supports a nice physical interpretation as a flux function. Suppose that we compute the magnetic flux through a spherical cap at a given radius r_0, from $\theta=0$ to $\theta=\Theta$. Details on the covariant definition of surface fluxes in general relativity are given in the Appendix 5.B. Using Eq. (5.39), the magnetic flux reads

$$\Phi=\int_{\theta=0}^{\Theta}\int_{\varphi=0}^{2\pi}B^r\sqrt{h}\,\mathrm{d}\theta\,\mathrm{d}\varphi=2\pi\int_{\theta=0}^{\Theta}\partial_\theta A_\varphi\,\mathrm{d}\theta=2\pi\left(A_\varphi(r_0,\theta=\Theta)-A_\varphi(r_0,\theta=0)\right). \tag{5.40}$$

Since the 4-potential is defined up to a constant value, most of the time one can choose $A_\varphi(r,0)=0$ without loss of generality. Up to a factor 2π, A_φ is interpreted as the magnetic flux through this spherical cap.

Another convenient property of A_φ is that it remains constant along poloidal magnetic field lines (it is also sometimes called stream function), if the fields are axisymmetric. To show this, one simply needs to check that $\boldsymbol{B}_\mathrm{p} \cdot \boldsymbol{\nabla} A_\varphi = 0$, with $\boldsymbol{B}_\mathrm{p}$ the poloidal magnetic field. This is true because $B^r = \partial_\theta A_\varphi / \sqrt{h}$ and $B^\theta = -\partial_r A_\varphi / \sqrt{h}$.

5.3.2 Interpretation

Equations (5.37a), (5.37b), (5.37d) and (5.37c) are identical in form to Maxwell's equations in matter (Bertin et al., 1996). In that case, the field $\boldsymbol{D}$ is the electric displacement which includes a contribution from the density of bound charges due to the polarization of matter, and $\boldsymbol{H}$ is the magnetizing field, which includes a contribution of the magnetization of matter. If the medium is linear, isotropic and instantly-responding, there can exist proportionality relations between $\boldsymbol{D} = \epsilon \boldsymbol{E}$ (ϵ being the electric permittivity of the material), and $\boldsymbol{B} = \mu \boldsymbol{H}$ (μ being the magnetic permeability of the material) between these fields. These relations are the constitutive relations of the medium.

We can draw an analogy with electromagnetism in curved spacetime by deriving similar constitutive relations. We can prove, using the definitions of $\boldsymbol{E}$, $\boldsymbol{D}$, $\boldsymbol{B}$ and $\boldsymbol{H}$ through the Maxwell tensor, that these relations are

$$\begin{cases} \boldsymbol{E} = \alpha \boldsymbol{D} + \boldsymbol{\beta} \times \boldsymbol{B}, & \text{(5.41a)} \\ \boldsymbol{H} = \alpha \boldsymbol{B} - \boldsymbol{\beta} \times \boldsymbol{D}. & \text{(5.41b)} \end{cases}$$

Proof

We will simply prove Eq. (5.41a), the proof of Eq. (5.41b) being very similar. Using Eq. (5.18) and the definitions $D^i = \alpha F^{0i}$, $E_i = F_{i0}$, we find that

$$\begin{aligned} D^i &= \alpha g^{0\mu} g^{i\nu} F_{\mu\nu} = -n^\mu g^{i\nu} F_{\mu\nu} && \text{(5.42)} \\ &= \frac{1}{\alpha} \left(-g^{i\nu} F_{0\nu} + g^{i\nu} \beta^j F_{j\nu} \right) && \text{(5.43)} \\ &= \frac{1}{\alpha} \left(-g^{ij} F_{0j} + g^{i0} \beta^j F_{j0} + g^{ik} \beta^j F_{jk} \right) && \text{(5.44)} \\ &= \frac{1}{\alpha} \left(-h^{ij} F_{0j} + \frac{\beta^i \beta^j}{\alpha^2} \left(F_{0j} + F_{j0} \right) + \beta^j h^{ik} F_{jk} - \frac{\beta^i}{\alpha^2} \beta^j \beta^k F_{jk} \right) && \text{(5.45)} \\ &= \frac{1}{\alpha} \left(E^i - \beta^j h^{ik} F_{kj} \right). && \text{(5.46)} \end{aligned}$$

In the last equality, the second and fourth terms are zero due to the antisymmetry of $F_{\mu\nu}$. Finally, using the Levi-Civita contraction (5.72), we note that

$$\begin{aligned} \left(\boldsymbol{\beta} \times \boldsymbol{B} \right)_i &= \sqrt{h} \varepsilon_{ijk} \beta^j B^k && \text{(5.47)} \\ &= \frac{1}{2} \beta^j \varepsilon_{ijk} \varepsilon^{kmn} F_{mn} && \text{(5.48)} \\ &= \frac{1}{2} \beta^j \left(\delta_i{}^m \delta_j{}^n - \delta_i{}^n \delta_j{}^m \right) F_{mn} = \beta^j F_{ij}; && \text{(5.49)} \end{aligned}$$

this is precisely the covariant component of the last term in Eq. (5.46). We conclude that $\alpha D^i = E^i - \left(\boldsymbol{\beta} \times \boldsymbol{B} \right)^i$.

Eq. (5.41a) and (5.41b) can be interpreted as the instantaneous response of curved spacetime, which differs from pure vacuum. For instance, an electric field can be "inductively" generated by the cross-product of the shift vector $\boldsymbol{\beta}$ and the magnetic field $\boldsymbol{B}$.

There remains to give a physical interpretation of these fields. Let us define the 4-vectors $\mathcal{B}^\mu = -*F^{\mu\nu}n_\nu$ and $\mathcal{D}^\mu = F^{\mu\nu}n_\nu$. These vectors verify $\mathcal{B}^t = 0$, $\mathcal{D}^t = 0$ and $\mathcal{B}^i = B^i$, $\mathcal{D}^i = D^i$, using Eq. (5.10). These 4-vectors are tangent to hypersurfaces Σ and reduce to the spatial 3-vectors $\boldsymbol{B}$ and $\boldsymbol{D}$. Because their definition matches the ones of the electric and magnetic fields in special relativity from the Maxwell tensor, as measured by an observer with velocity n^μ, we conclude that $\boldsymbol{B}$ and $\boldsymbol{D}$ are respectively the magnetic field and electric field measured by FIDOs (Thorne & MacDonald, 1982). They have a clear physical meaning. On the other hand, $\boldsymbol{H}$ and $\boldsymbol{E}$ are simply auxiliary fields, which can be interpreted as the magnetic field and electric field measured on the grid.

Similarly, the charge density is related to the 4-vector I^μ by $\rho = -I^\mu n_\mu$, so that ρ is also the charge density as measured by FIDOs. However, the 3-vector $\boldsymbol{J}$ cannot be identified with the current density measured by a FIDO. Actually, the 3-current density measured by a FIDO j^μ is the projection of I^μ on a hypersurface orthogonal to n^μ: $j^\mu = I^\mu + (n_\nu I^\nu)n^\mu$. This yields

$$j^i = I^i - \frac{\beta^i}{\alpha} n_\nu I^\nu \Rightarrow \boldsymbol{j} = \frac{\boldsymbol{J} + c\rho\boldsymbol{\beta}}{\alpha}. \tag{5.50}$$

The second term accounts for the fact that the coordinate grid is sliding with respect to FIDOs.

> The difference between $\boldsymbol{j}$ and $\boldsymbol{J}$ reflects the difference between the velocity of a particle as measured by a FIDO $\boldsymbol{V}$, and as measured with respect to the coordinate grid $\boldsymbol{v}$. We will show in Sec. 10.1.1 that for a particle with charge q and position $\boldsymbol{x}_0$, the current density that this particle generates is
>
> $$\boldsymbol{J} = q\boldsymbol{v}\delta\left(\boldsymbol{x} - \boldsymbol{x}_0\right), \; \boldsymbol{j} = q\boldsymbol{V}\delta\left(\boldsymbol{x} - \boldsymbol{x}_0\right). \tag{5.51}$$

5.3.3 Energetics

Finally, we need to evaluate the energy carried by the electromagnetic field. The energy-momentum tensor of the sole electromagnetic field, obtained via the minimal coupling procedure, reads

$$T^\mu{}_\nu = \frac{1}{4\pi}\left(F^{\mu\alpha}F_{\nu\alpha} - \frac{1}{4}\left(F_{\alpha\beta}F^{\alpha\beta}\right)\delta^\mu{}_\nu\right). \tag{5.52}$$

This tensor is constructed to verify the conservation of energy-momentum:

$$\nabla_\mu T^\mu{}_\nu = -F_{\nu\alpha}I^\alpha. \tag{5.53}$$

In this case, it is a little trickier to derive a transparent conservation equation, but this still works nicely for the time part, which is the energy conservation law. By performing a $3+1$ splitting of Eq. (5.52) we obtain the following equation:

$$\partial_t e + \boldsymbol{\nabla}\cdot\boldsymbol{\Pi} = -\boldsymbol{E}\cdot\boldsymbol{J}, \tag{5.54}$$

where

$$e = -\alpha T^0{}_0 = \frac{1}{8\pi}\left(\boldsymbol{E}\cdot\boldsymbol{D} + \boldsymbol{B}\cdot\boldsymbol{H}\right), \; \Pi^i = -c\alpha T^i{}_0 = \frac{c}{4\pi}\left(\boldsymbol{E}\times\boldsymbol{H}\right)^i. \tag{5.55}$$

Proof

The covariant derivative in Eq. (5.52) can be expanded using the Christoffel symbols, using again Eq. (5.80):

$$\nabla_\mu T^\mu{}_\nu = \partial_\mu T^\mu{}_\nu + \Gamma^\mu{}_{\mu\lambda} T^\lambda{}_\nu - \Gamma^\lambda{}_{\nu\mu} T^\mu{}_\lambda = \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} T^\mu{}_\nu \right) - \Gamma_{\lambda\mu\nu} T^{\lambda\mu}, \tag{5.56}$$

where we have used the symmetry of the energy-momentum tensor $T^{\lambda\mu} = T^{\mu\lambda}$ and the symmetry of the Christoffel symbols $\Gamma_{\lambda\mu\nu} = \Gamma_{\lambda\nu\mu}$. Unlike the $3+1$ splitting of Maxwell's equations, here the second term does not vanish *a priori*. We use the fact that the Christoffel symbols verify the identity $\Gamma_{\lambda\mu\nu} + \Gamma_{\mu\lambda\nu} = \partial_\nu g_{\mu\lambda}$ to get

$$\Gamma_{\lambda\mu\nu} T^{\lambda\mu} = \frac{1}{2} \left(\Gamma_{\lambda\mu\nu} + \Gamma_{\mu\lambda\nu} \right) T^{\lambda\mu} = \frac{1}{2} T^{\lambda\mu} \partial_\nu g_{\lambda\mu}. \tag{5.57}$$

Using Eq. (5.19), this allows us to rewrite Eq. (5.53) as

$$\frac{1}{\sqrt{h}} \partial_\mu \left(\alpha \sqrt{h} T^\mu{}_\nu \right) = \frac{1}{2} \alpha T^{\lambda\mu} \partial_\nu g_{\lambda\mu} - \alpha F_{\mu\nu} I^\nu. \tag{5.58}$$

Now in the special case of a stationary metric, we have $\partial_t g_{\mu\nu} = 0$, so that in the time part of this equation the first term on the right-hand side vanishes. This way, we indeed obtain Eq. (5.54) by defining $e = -\alpha T^0{}_0$ and $\Pi^i = -c\alpha T^i{}_0$.

Eq. (5.54) is the $3+1$ equivalent of Poynting's theorem, where e is the volume density of electromagnetic energy-at-infinity and $\boldsymbol{\Pi}$ is the surface flux of energy-at-infinity (*i.e.* the Poynting flux). Note that these quantities are not measured by FIDOs.

Finally, let us stress that the Maxwell tensor is associated with Lorentz invariants $F_{\mu\nu}F^{\mu\nu}$ and $*F_{\mu\nu}F^{\mu\nu}$. The latter, in particular, can be rewritten as a function of the 3-spatial fields:

$$*F_{\mu\nu}F^{\mu\nu} = \frac{1}{2} e^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} \tag{5.59}$$

$$= \frac{1}{2\sqrt{-g}} 4\varepsilon^{0\nu\alpha\beta} F_{0\nu} F_{\alpha\beta} \tag{5.60}$$

$$= -\frac{4}{2\alpha\sqrt{h}} \varepsilon^{ijk} F_{0i} F_{jk} \tag{5.61}$$

$$= \frac{4}{\alpha} F_{i0} \frac{1}{2} e^{ijk} F_{jk} = 4 \frac{1}{\alpha} E_i B^i, \tag{5.62}$$

where we used $E_i = F_{i0}$, $B^i = (1/2) e^{ijk} F_{jk}$ and $\varepsilon^{0123} = -1 = -\varepsilon_{0123}$. Finally, using Eq. (5.41a), we obtain

$$*F_{\mu\nu}F^{\mu\nu} = 4\boldsymbol{D} \cdot \boldsymbol{B}. \tag{5.63}$$

The second Lorentz invariant has a clear interpretation as the electric field parallel to the magnetic field lines (the "parallel electric field" in the following), as measured by FIDOs. Similarly, the other Lorentz invariant is $F_{\mu\nu}F^{\mu\nu} = 2\left(\boldsymbol{B}^2 - \boldsymbol{D}^2\right)$. These invariants have the same expression in flat spacetime.

5.A Appendix: Useful notions on tensor analysis

5.A.1 Levi-Civita tensor

This section relies heavily on Blau (2018) and Gourgoulhon (2010).

Definition

The Levi-Civita symbol $\varepsilon_{\mu\nu\rho\sigma}$ is totally antisymmetric: it takes value 0 if two of the four indices (μ,ν,ρ,σ) are equal, 1 if (μ,ν,ρ,σ) can be deduced from $(0,1,2,3)$ by an even permutation, and -1 otherwise. It is merely a symbol, and has the same components in all coordinate systems. For this reason, it is not a tensor.

Consider a direct orthonormal basis of the spacetime $(\boldsymbol{e}_i^*)$, associated with an inertial coordinate system (x^i). In 4 dimensions, there is but one 4-form e which is totally antisymmetric and verifies

$$e\left(\boldsymbol{e}_0^*,\boldsymbol{e}_1^*,\boldsymbol{e}_2^*,\boldsymbol{e}_3^*\right)=+1. \tag{5.64}$$

The 4-form e is the Levi-Civita *tensor*. Let us determine its components in any coordinate system (y^i), associated with the natural basis $(\boldsymbol{e}_i)$. The change-of-basis matrix from $(\boldsymbol{e}_i^*)$ to $(\boldsymbol{e}_i)$ is the Jacobi matrix with elements $J^\mu{}_\nu=\partial x^\mu/\partial y^\nu$, so that $\boldsymbol{e}_\nu=J^\mu{}_\nu\boldsymbol{e}_\mu^*$. Now, because e is a 4-linear form we have

$$\begin{aligned} e(\boldsymbol{e}_0,\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3) &= J^\mu{}_0\,J^\nu{}_1\,J^\alpha{}_2\,J^\beta{}_3\,e(\boldsymbol{e}_0^*,\boldsymbol{e}_1^*,\boldsymbol{e}_2^*,\boldsymbol{e}_3^*) &(5.65)\\ &= \sum_\sigma (-1)^k(\sigma)J^{\sigma(0)}{}_0\,J^{\sigma(1)}{}_1\,J^{\sigma(2)}{}_2\,J^{\sigma(3)}{}_3=\det J, &(5.66)\end{aligned}$$

where the sum is on all permutations of 4 elements, and $k(\sigma)$ is the signature of the permutation σ. To conclude, we simply need to relate the determinant of the Jacobi matrix with that of the metric. A component of the metric can be written as

$$g_{\alpha\beta}=\boldsymbol{e}_\alpha\cdot\boldsymbol{e}_\beta=J^\mu{}_\alpha\,J^\mu{}_\beta\,\boldsymbol{e}_\mu^*\cdot\boldsymbol{e}_\nu^*. \tag{5.67}$$

Because $(\boldsymbol{e}_i^*)$ is an orthonormal basis, it is associated with the Minkowski matrix, with determinant -1. Therefore we deduce $g=-\left(\det J\right)^2\Rightarrow e(\boldsymbol{e}_0,\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3)=\sqrt{-g}$. This also shows that for any coordinate systems, g is negative. All in all, the Levi-Civita tensor has components

$$e_{\mu\nu\rho\sigma}=\sqrt{-g}\varepsilon_{\mu\nu\rho\sigma}. \tag{5.68}$$

> From what we have already shown, it is straightforward to prove that the Levi-Civita tensor defined by Eq. (5.68) indeed transforms as a tensor under a change of coordinates $(x^\mu)\to(y^\nu)$, that is,
>
> $$e_{\mu'\nu'\alpha'\beta'}=J^\mu{}_{\mu'}\,J^\nu{}_{\nu'}\,J^\alpha{}_{\alpha'}\,J^\beta{}_{\beta'}\,e_{\mu\nu\alpha\beta} \tag{5.69}$$

Likewise, it is possible to define the totally antisymmetric $(4,0)$ tensor $e^{\mu\nu\rho\sigma}=\varepsilon^{\mu\nu\rho\sigma}/\sqrt{-g}$, with $\varepsilon^{\mu\nu\rho\sigma}$ the totally antisymmetric symbol with $\varepsilon^{0123}=-1=-\varepsilon_{0123}$. All the notions have been introduced in a 4-dimensional spacetime, but are also valid on a 3-dimensional manifold. In particular, a Levi-Civita tensor can also be defined on a particular spacelike hypersurface of a foliation (see Sec. 5.1.2), where the determinant of the induced spatial metric $h>0$ should be used instead of $-g$. Incidentally, we have also shown that upon such a

coordinate change with Jacobi matrix J, the determinant of the matrix transforms as

$$\sqrt{-g'} = |\det J|\sqrt{-g}. \tag{5.70}$$

This will be useful when trying to define generally covariant integration in Sec. 5.B.

Curl of a vector

The Levi-Civita tensor can be used to define the curl $\boldsymbol{\nabla} \times \boldsymbol{V}$ of a 3-vector $\boldsymbol{V}$:

$$(\boldsymbol{\nabla} \times \boldsymbol{V})^i = e^{ijk}\partial_j V_k = e^{ijk}\nabla_j V_k = \frac{1}{\sqrt{h}}\varepsilon^{ijk}\partial_j V_k. \tag{5.71}$$

The second equation results from the symmetry of the Christoffel symbols and the antisymmetry of the Levi-Civita tensor.

Useful contractions

Finally, let us state two useful contractions of the Levi-Civita symbol in 3 dimensions:

$$\varepsilon_{jki}\varepsilon^{imn} = \delta_j{}^m\delta_k{}^n - \delta_j{}^n\delta_k{}^m \tag{5.72}$$

$$\varepsilon_{kij}\varepsilon^{ijn} = 2\delta_k{}^n. \tag{5.73}$$

5.A.2 Covariant divergence

The divergence of a 4-vector V_μ can be defined as $\nabla_\mu V^\mu = \partial_\mu V^\mu + \Gamma^\mu{}_{\mu\nu}V^\nu$, where $\Gamma^\alpha{}_{\beta\gamma}$ is the Christoffel symbol associated with the metric. One can show the following useful identity:

$$\Gamma^\mu{}_{\mu\lambda} = \frac{1}{\sqrt{-g}}\partial_\lambda\sqrt{-g}. \tag{5.74}$$

Proof

The contracted Christoffel symbol can be expressed as

$$\Gamma^\mu{}_{\mu\lambda} = \frac{1}{2}g^{\mu\nu}\partial_\lambda g_{\mu\nu}. \tag{5.75}$$

The determinant of the metric can be written as

$$g = \frac{1}{4!}\varepsilon^{\alpha\beta\gamma\delta}\varepsilon^{\mu\nu\rho\sigma}g_{\alpha\mu}g_{\beta\nu}g_{\gamma\rho}g_{\delta\sigma}. \tag{5.76}$$

In this equation, $\varepsilon_{\mu\nu\rho\sigma}$ is the Levi-Civita *symbol*. Therefore, a small variation of g, viewed as a smooth function of the coefficients $g_{\mu\nu}$, reads

$$\delta g = \frac{1}{3!}\varepsilon^{\alpha\beta\gamma\delta}\varepsilon^{\mu\nu\rho\sigma}g_{\alpha\mu}g_{\beta\nu}g_{\gamma\rho}\delta g_{\delta\sigma} \tag{5.77}$$

$$= 4\left(\frac{1}{4!}\varepsilon^{\alpha\beta\gamma\delta}\varepsilon^{\mu\nu\rho\sigma}g_{\alpha\mu}g_{\beta\nu}g_{\gamma\rho}g_{\delta\sigma}\right)\frac{g^{\delta\sigma}}{4}\delta g_{\delta\sigma} \tag{5.78}$$

$$= gg^{\delta\sigma}\delta g_{\delta\sigma}, \tag{5.79}$$

where we used $g^{\mu\nu}g_{\mu\nu} = 4$. It follows that $\partial_\lambda g = gg^{\mu\nu}\partial_\lambda g_{\mu\nu}$, and then $\partial_\lambda\sqrt{-g} = (1/2)\sqrt{-g}g^{\mu\nu}\partial_\lambda g_{\mu\nu}$. Inserting this into Eq. (5.75) yields the result of Eq. (5.74).

Thus, the covariant divergence can be rewritten as

$$\nabla_\mu V^\mu = \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}V^\mu\right). \tag{5.80}$$

This identity is part of the reason why the determinant of the metric is an important quantity; it is the only quantity needed to compute the divergence of a vector field. This formula is useful to find the expression of the divergence in various coordinate systems (cylindrical, spherical), even in flat space.

This result can also be used to simplify the expression of the divergence of an antisymmetric tensor. For a $(2,0)$ antisymmetric tensor $T^{\mu\nu}$ for instance, we have

$$\nabla_\mu T^{\mu\nu} = \partial_\mu T^{\mu\nu} + \Gamma^\mu{}_{\mu\lambda}T^{\lambda\nu} + \Gamma^\nu{}_{\mu\lambda}T^{\mu\lambda} = \frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}T^{\mu\nu}\right), \tag{5.81}$$

because in the last term on the right-hand side, the contraction of the symmetric Christoffel symbol with the antisymmetric tensor $T^{\mu\lambda}$ is zero.

5.B Appendix: Covariant integration

The goal is to find a generally covariant definition of the integration of a scalar field over some volume, or the flux of a vector field across a surface. The first thing to notice is that the infinitesimal volume element is not itself invariant under a coordinate transformation: for a transformation $(x^\mu) \to (y^\nu)$ parametrized by the Jacobian matrix J, we have

$$\mathrm{d}^4y = \left|\det\left(\frac{\partial y^\nu}{\partial x^\mu}\right)\right|\mathrm{d}^4x = \frac{1}{|\det J|}\,\mathrm{d}^4x. \tag{5.82}$$

However, we noted in Sec. 5.A.1 that the determinant transforms as $\sqrt{-g'} = |\det J|\sqrt{-g}$ under such a transformation. Therefore, the way to measure a volume in a coordinate-independent manner is to define the volume element as $\sqrt{-g}\,\mathrm{d}^4x$, which is invariant under a general coordinate transformation. For a given volume $\mathcal{V}$ in spacetime, we define the integration of a scalar field f over $\mathcal{V}$ as

$$\int_{\mathcal{V}} f = \int_{\mathcal{V}} f(x^\mu)\sqrt{-g}\,\mathrm{d}^4x. \tag{5.83}$$

Similarly, it is possible to perform the integration of a scalar over a spacelike hypersurface Σ. If h_{ij} is the metric on Σ induced by the metric of the spacetime $g_{\mu\nu}$ and y^i is a coordinate system on Σ, one must simply replace $\sqrt{-g}\,\mathrm{d}^4x$ by $\sqrt{h}\,\mathrm{d}^3y$ in Eq. (5.83).

Now, we might want to compute integrals of other quantities than just scalar fields; for exemple the flux of the magnetic field through the event horizon of the black hole, or through any sphere of constant radius r. The goal of this section will be to provide a way to compute such fluxes in a covariant way. Let us denote $\mathcal{S}$ the 2-dimensional boundary hypersurface of a volume V, which we assume spacelike. $\mathcal{S}$ is endowed with the 2-dimensional induced metric ρ_{ab}, and coordinates ξ^i. Let J^μ be a vector field. We can make use of the identity (5.80) to derive the relativistic analog of Gauss' theorem. Using the definition (5.83) one has

$$\int_V \nabla_\mu J^\mu\sqrt{h}\,\mathrm{d}^3y = \int_V \partial_\mu\left(\sqrt{h}J^\mu\right)\mathrm{d}^3y. \tag{5.84}$$

Now the right-hand side is the integral of an ordinary total derivative.

We assume that we are working with an adapted coordinate system $y^i = (S, y^j)$, such that the boundary $\mathcal{S}$ has the equation $S = \mathrm{cst}$. In particular, it is so for our case of interest: a

sphere of constant radius $r = R$. In the following we will assume $S = r$, but one can choose an adapted coordinate system for any surface. Thus we have

$$\int_V \nabla_\mu J^\mu \sqrt{h}\, \mathrm{d}^3 y = \int_{\mathcal{S}} J^r \sqrt{h}\, \mathrm{d}^2\xi. \tag{5.85}$$

Eq. (5.85) shows that the quantity $\int_{\mathcal{S}} J^r \sqrt{h}\, \mathrm{d}^2\xi$ appearing on the right-hand side is a generally covariant definition of a flux integral over a surface of constant r. This also provides a convenient way to compute such a flux, using this generalization of Gauss' theorem, if the divergence of the vector field is known. It is not *a priori* obvious that the $\sqrt{h}$ factor should be in this definition, instead of the determinant ρ of the 2-dimensional metric on $\mathcal{S}$. This expression is the most convenient, and we will use it extensively.

In fact, it is possible to make a connection with the 2-metric ρ_{ab}, which allows to define an oriented surface element on Σ. One can define a vector normal to the surface $\mathcal{S}$, defined by $r = \mathrm{cst}$, by noting that a vector field tangent to $\mathcal{S}$ must verify $V^i \partial_i r = 0$ (r does not vary along directions tangent to $\mathcal{S}$). Therefore, a normal vector $\boldsymbol{N}$ to $\mathcal{S}$ has components

$$N_i = \frac{\partial_i r}{\sqrt{h^{ij}\partial_i r \partial_j r}} = \frac{\partial_i r}{\sqrt{h^{rr}}}, \quad N^i = \frac{h^{ir}}{\sqrt{h^{rr}}}. \tag{5.86}$$

In particular, we have $N_r = 1/\sqrt{h^{rr}}$. In addition to that, one can use Cramer's rule just like in Sec. 5.1.4 to express h^{rr} as $h^{rr} = \rho/h$. We conclude that

$$\begin{aligned}
\int_V \nabla_\mu J^\mu \sqrt{h}\, \mathrm{d}^3 y &= \int_{\mathcal{S}} J^r \sqrt{h}\, \mathrm{d}^2\xi && (5.87)\\
&= \int_{\mathcal{S}} \frac{h^{ir}}{\sqrt{h^{rr}}} J_i \sqrt{\rho}\, \mathrm{d}^2\xi && (5.88)\\
&= \int_{\mathcal{S}} N^i J_i \sqrt{\rho}\, \mathrm{d}^2\xi = \int_{\mathcal{S}} N_i J^i \sqrt{\rho}\, \mathrm{d}^2\xi. && (5.89)
\end{aligned}$$

This is the most proper version of Gauss' theorem, as it equates a manifestly covariant volume integral and a manifestly covariant surface integral. The oriented surface element of $\mathcal{S}$ can be defined as

$$\mathrm{d}S^i = \sqrt{\rho} N^i\, \mathrm{d}^2\xi. \tag{5.90}$$

Example

As an example, we will use these notions to compute the surface area of a Kerr black hole (Bicak & Janis, 1985). In the Boyer-Lindquist coordinates, the metric ρ_{ab} induced on the 2-sphere, spanned by the coordinates (θ, φ) and defined by $t = \mathrm{cst}$ and $r = r_\mathrm{h}$, is given by

$$\mathrm{d}s^2 = \Sigma\, \mathrm{d}\theta^2 + \frac{A \sin^2\theta}{\Sigma}\, \mathrm{d}\varphi^2. \tag{5.91}$$

The determinant of the metric is $\rho = \rho_{\theta\theta}\rho_{\varphi\varphi}$. At the event horizon, we have $\Delta(r_\mathrm{h}) = 0$, so that $\rho = \left(r_\mathrm{h}^2 + a^2 r_\mathrm{g}^2\right)^2 \sin^2\theta$. The total surface area of the event horizon is

$$S = \int_0^{2\pi}\int_0^{\pi} \sqrt{\rho}\, \mathrm{d}\theta\, \mathrm{d}\varphi = 4\pi\left(r_\mathrm{h}^2 + a^2 r_\mathrm{g}^2\right). \tag{5.92}$$

Chapter 6

Black-hole electrodynamics

Now that we have formulated electrodynamics in the language of general relativity, our goal will be to derive several solutions of Maxwell's equations around a black hole. This is no easy task, and we will be able to find only a few solutions. In Sec. 6.1, we will first derive vacuum electromagnetic fields, which will give us insight to study the main topic of interest: how such fields interact with plasma. In Sec. 6.2 we will see how magnetospheres are modified by the presence of plasma in the force-free limit. The present chapter sets the stage for Chap. 7, in which we will describe the Blandford-Znajek solution in the force-free limit. In all this work, we neglect the backreaction of the energy and momentum of the electromagnetic field on the metric itself.

6.1 Vacuum solutions

6.1.1 Killing vectors and Maxwell's equations

Our first goal is to construct solutions of Maxwell's equations in vacuum, by deriving the 4-potential A_μ. Actually, such vacuum solutions can be constructed from the Killing vectors of the Kerr spacetime. Let us first prove a useful identity relating the derivatives of a Killing vector K^μ to the curvature Riemann tensor $R^\rho{}_{\lambda\mu\nu}$ (Blau, 2018):

$$\nabla_\mu \nabla_\nu K_\rho = R^\sigma{}_{\mu\nu\rho} K_\sigma. \tag{6.1}$$

Proof

By definition of the Riemann tensor, we have $[\nabla_\mu, \nabla_\nu] V_\lambda = -R^\rho{}_{\lambda\mu\nu} V_\rho$ for any 1-form V_μ (noting $[\nabla_\mu, \nabla_\nu] = \nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu$). A Killing vector is defined by Eq. (4.16), so we get $\nabla_\mu \nabla_\nu K_\rho + \nabla_\mu \nabla_\rho K_\nu = 0$. Combinations of this identity obtained by switching indices yield:

$$\begin{aligned} 0 = &(\nabla_\mu \nabla_\nu K_\rho + \nabla_\mu \nabla_\rho K_\nu) - (\nabla_\rho \nabla_\mu K_\nu + \nabla_\rho \nabla_\nu K_\mu) \\ &+ (\nabla_\nu \nabla_\rho K_\mu + \nabla_\nu \nabla_\mu K_\rho) - (\nabla_\mu \nabla_\nu K_\rho + \nabla_\mu \nabla_\rho K_\nu). \end{aligned} \tag{6.2}$$

This can be rewritten as

$$0 = \nabla_\mu \nabla_\nu K_\rho - \nabla_\mu \nabla_\rho K_\nu + [\nabla_\mu, \nabla_\rho] K_\nu + [\nabla_\nu, \nabla_\rho] K_\mu + [\nabla_\nu, \nabla_\mu] K_\rho. \tag{6.3}$$

The first two terms are identical because K is a Killing vector. We finally obtain

$$\nabla_\mu \nabla_\nu K_\rho = \frac{1}{2}(R_{\sigma\nu\mu\rho} + R_{\sigma\mu\nu\rho} + R_{\sigma\rho\nu\mu}) V^\sigma = R_{\sigma\mu\nu\rho} K^\sigma = R^\sigma{}_{\mu\nu\rho} K_\sigma, \tag{6.4}$$

where we used the cyclic permutation symmetry of the Riemann tensor: $R_{\alpha[\beta\gamma\delta]} = 0$.

Now we can contract Eq. (6.1) over μ and ρ to get $\nabla^\rho\nabla_\nu K_\rho = R^\mu{}_\nu K_\mu$, with $R^\mu{}_\nu$ the Ricci tensor. Using again the definition $\nabla_\mu K_\nu + \nabla_\nu K_\mu = 0$, we see that Killing vectors obey the two following equations:

$$\nabla_\mu K^\mu = 0; \quad \nabla_\nu\nabla^\nu K^\mu = -R^\mu{}_\nu K^\nu. \tag{6.5}$$

On the other hand, if A_μ is a solution of the vacuum Maxwell's equations (with $I^\nu = 0$), we know from Eq. (5.27) that it must verify, in the Lorenz gauge,

$$\nabla_\mu A^\mu = 0; \quad \nabla_\nu\nabla^\nu A^\mu = R^\mu{}_\nu A^\nu. \tag{6.6}$$

The sign of the Ricci tensor is different in Eq. (6.5) and (6.6), but this difference disappears with vanishing Ricci tensor, that is, outside of any matter source. In that case, K^μ and A^μ verify the exact same equations. Therefore, any Killing vector is a solution of the vacuum Maxwell's equations in a gravitational field, by identifying the potential vector A^μ with K^μ.

Example

The resulting electromagnetic field may sometimes be trivial. In flat spacetime, a general Killing vector has the form $K^\alpha = \omega^\alpha_\beta x^\beta + b^\alpha$, with $\omega_{\alpha\beta} = -\omega_{\beta\alpha}$ (which can readily be seen from solving $\partial_a K_b + \partial_b K_a = 0$). The associated Maxwell tensor is $F_{\alpha\beta} = \partial_\alpha K_\beta - \partial_\beta K_\alpha = -2\omega_{\alpha\beta}$. The four translational Killing vectors b^α generate no electromagnetic field. The three spatial rotations generate the three components of a uniform magnetic field, whereas the three boosts generate the three components of a uniform electric field.

As said earlier, the Kerr spacetime possesses a timelike Killing vector $\boldsymbol{\eta} = \partial_t$ and an axial Killing vector $\boldsymbol{\psi} = \partial_\varphi$.

6.1.2 Multipoles

Electric monopole

Let us first investigate the electromagnetic field generated by $\boldsymbol{\eta}$, in BL coordinates. Its covariant components are

$$\eta_\mu = (g_{tt}, 0, 0, g_{t\varphi}) = \left(\frac{2r_\mathrm{g} r}{\Sigma} - 1, 0, 0, -\frac{2ar_\mathrm{g}^2 r\sin^2\theta}{\Sigma}\right). \tag{6.7}$$

To describe the qualitative properties of the field, let us look at the Schwarzschild limit $a = 0$: the components η_μ reduce to $\eta_\mu = (2r_\mathrm{g}/r - 1, 0, 0, 0)$. The associated "electric field" is $E_i = \partial_i\eta_t = -\left(2r_\mathrm{g}/r^2\right)\delta^r{}_i$. We see that this electric field is that of a point charge at $r = 0$, $-\eta_t$ being the "electric potential". The timelike Killing vector $\boldsymbol{\eta}$ generates the field of an electric monopole.

Thus, we can make the guess that the 4-potential A_η, associated with the Maxwell tensor $F_{\eta_{\mu\nu}}$, with components

$$A_{\eta_\mu} = -\frac{Q}{2r_\mathrm{g}}\eta_\mu = -\frac{Q}{2r_\mathrm{g}}\left(\frac{2r_\mathrm{g} r}{\Sigma} - 1, 0, 0, -\frac{2ar_\mathrm{g}^2 r\sin^2\theta}{\Sigma}\right), \tag{6.8}$$

describes an electric monopole supported by a Kerr black hole with charge Q (the constant -1 term is irrelevant and could be dropped). To check this, we need to unambiguously compute the charge carried by the black hole, and not just rely on the Schwarzschild limit.

Let us define a closed 2-sphere $\mathcal{S}_r$ of radius r, at constant t, centered on $r = 0$, enclosing the volume $\mathcal{V}_r$. The electric charge Q_η enclosed by this sphere can be computed by Gauss' theorem. Using Eq. (5.37a), Eq. (5.85) and the definition $D^i = \alpha F^{ti}$, we have

$$4\pi Q_\eta = \int_{\mathcal{V}_r} (\boldsymbol{\nabla} \cdot \boldsymbol{D}) \sqrt{h}\, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\varphi \tag{6.9}$$

$$= \int_{\mathcal{S}_r} D^r \sqrt{h}\, \mathrm{d}\theta\, \mathrm{d}\varphi \tag{6.10}$$

$$= \int_{\mathcal{S}_r} F_\eta^{tr} \sqrt{-g}\, \mathrm{d}\theta\, \mathrm{d}\varphi = \int_{\mathcal{S}_r} {*F}_{\eta\theta\varphi}\, \mathrm{d}\theta\, \mathrm{d}\varphi, \tag{6.11}$$

where we used the definition of the Hodge dual ${*F}_{\mu\nu} = (1/2) e_{\mu\nu\alpha\beta} F^{\alpha\beta} \Rightarrow {*F}_{\eta\theta\varphi} = \sqrt{-g} F^{tr}$. Although Eq. (6.11) is not manifestly covariant, it is indeed invariant under a change of coordinates due to F_η being antisymmetric.

> This integral is actually related to the *Komar mass*, which can be used to define the black hole's mass. Because the field strength F_η derives from a Killing vector, it is possible to show that this integral does not depend on the specific choice of the surface $\mathcal{S}_r$, as long as all matter is contained in this surface (Gourgoulhon, 2012). For this reason, the Komar mass is a global quantity, an intrinsic property of spacetime.

To simplify the calculation of this integral, we will suppose that the radius of the sphere is much larger than r_g, since the result should not depend on r. Using the expressions of the coefficients of the inverse metric and the determinant, the component ${*F}_{\eta\theta\varphi}$ reads

$${*F}_{\eta\theta\varphi} = \sqrt{-g} \left(g^{00} g^{rr} F_{0r} + g^{0\varphi} g^{rr} F_{\varphi r} \right) \tag{6.12}$$

$$= \Sigma \sin\theta \left(-\frac{A}{\Delta\Sigma} \frac{\Delta}{\Sigma} \left(-\partial_r A_{\eta 0} \right) - \frac{2 r_\mathrm{g}^2 r a}{\Delta\Sigma} \frac{\Delta}{\Sigma} \left(-\partial_r A_{\eta\varphi} \right) \right) \tag{6.13}$$

$$= \sin\theta \left(\frac{A}{\Sigma} \partial_r A_{\eta 0} + \frac{2 r_\mathrm{g}^2 r a}{\Sigma} \partial_r A_{\eta\varphi} \right). \tag{6.14}$$

In the limit $r \gg r_\mathrm{g}$, we have $A/\Sigma \simeq r^2$ and $\partial_r A_{\eta 0} \simeq Q/r^2$, whereas the second term on the right-hand side vanishes as $r \to +\infty$. We have finally

$$\int_{\mathcal{S}_r} {*F}_{\eta\theta\varphi}\, \mathrm{d}\theta\, \mathrm{d}\varphi = \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} Q \sin\theta\, \mathrm{d}\theta\, \mathrm{d}\varphi \Rightarrow Q_\eta = Q. \tag{6.15}$$

Similarly, the magnetic charge P_η associated with the 4-potential A_η can be computed. We use the magnetic version of Gauss' theorem, using $B^r = (1/2) e^{rjk} F_{\eta ij} = F_{\eta\theta\varphi}/\sqrt{h}$:

$$4\pi P_\eta = \int_{\mathcal{V}_r} \boldsymbol{\nabla} \cdot \boldsymbol{B} \sqrt{h}\, \mathrm{d}r\, \mathrm{d}\theta\, \mathrm{d}\varphi \tag{6.16}$$

$$= \int_{\mathcal{S}_r} B^r \sqrt{h}\, \mathrm{d}\theta\, \mathrm{d}\varphi \tag{6.17}$$

$$= \int_{\mathcal{S}_r} F_{\eta\theta\varphi}\, \mathrm{d}\theta\, \mathrm{d}\varphi. \tag{6.18}$$

This equality underlines the duality between electric and magnetic fields via the Hodge dual operation. We have

$$F_{\eta\theta\varphi} = \partial_\theta A_{\eta_\varphi} = -2Qar_{\rm g} r\cos\theta\sin\theta\frac{r^2+a^2r_{\rm g}^2}{\Sigma^2}. \tag{6.19}$$

The integral of $F_{\eta\theta\varphi}$ over θ vanishes for all r. Consequently, the field generated by A_η carries no magnetic charge. Now we know that the 4-potential A_η describes a stationary, axisymmetric field which has charge Q, no magnetic monopole and vanishes asymptotically. A special case of the no-hair theorem actually assures that this solution is unique (Ipser, 1971; Wald, 1972):

Theorem 6.1.1. *Let F be the Maxwell tensor of a stationary, axisymmetric, vacuum black hole space time. If F is stationary, axisymmetric and nonsingular at the event horizon, vanishes at large distances and has no electric charge or monopole moment, then $F = 0$.*

There can be at most one stationary, axisymmetric, vacuum solution of Maxwell's equations endowing the black hole with a charge Q and no magnetic charge. A_η is the 4-potential of an electrically charged Kerr-Newman black hole. The electric field reduces to that of a point charge if the spin parameter is zero. That being said, if $a \neq 0$, a dipolar magnetic field arises due to the dragging of inertial frames into rotation, just like a charged conducting sphere would have an induced magnetic moment if it were forced to spin.

Magnetic monopole

Altough magnetic charges most likely do not exist, it will come in handy to have an exact solution for a magnetic monopole in Kerr spacetime, as a test case for the code's electromagnetic solver. In order to construct the solution for a magnetic monopole, it seems logical to take the dual of F_η, which is also a solution of the vacuum Maxwell's equations (the Hodge dual of a 2-form is another 2-form). The calculation of the different components of $*F_\eta$ are lengthy and not particularly illuminating. We will simply give the expressions of the non-vanishing components, where the electric charge Q is replaced by a magnetic charge P:

$$*F_{\eta\theta\varphi} = \sqrt{-g}\,F_\eta^{0r} = P\frac{\sin\theta}{\Sigma^2}\left(r^2+a^2r_{\rm g}^2\right)\left(r^2-a^2r_{\rm g}^2\cos^2\theta\right) \tag{6.20}$$

$$*F_{\eta 0r} = \sqrt{-g}\,F^{\theta\varphi} = P\frac{2ar_{\rm g} r\cos\theta}{\Sigma^2} \tag{6.21}$$

$$*F_{\eta 0\theta} = -\sqrt{-g}\,F^{r\varphi} = -P\frac{ar_{\rm g}\sin\theta}{\Sigma^2}\left(r^2-a^2r_{\rm g}^2\cos^2\theta\right) \tag{6.22}$$

$$*F_{\eta r\varphi} = -\sqrt{-g}\,F^{0\theta} = P\frac{2a^2r_{\rm g}^2 r\cos\theta\sin^2\theta}{\Sigma^2} \tag{6.23}$$

Since this is a solution of the vacuum Maxwell's equations, $*F_\eta$ must derive from a 4-potential $\overline{A}_\mu$. Integrating the components of $*F_\eta$ in r and θ, one can find that it derives from the following 4-potential,

$$\overline{A}_\mu = P\left(\frac{ar_{\rm g}\cos\theta}{\Sigma}, 0, 0, -\cos\theta\frac{r^2+a^2r_{\rm g}^2}{\Sigma}\right). \tag{6.24}$$

To interpret this new solution, let us look at the Schwarzschild limit $a = 0$: the 4-potential reduces to $\overline{A}_\mu = P\,(0,0,0,-\cos\theta)$. In this case, up to an additive constant, $A_\varphi = P\,(1-\cos\theta)$ is indeed the flux function of a monopole with charge P. Another way to realize this is to compute the components of the magnetic field: $B^r = \partial_\theta A_\varphi/\left(r^2\sin\theta\right) = P/r^2$, $B^\theta = 0$,

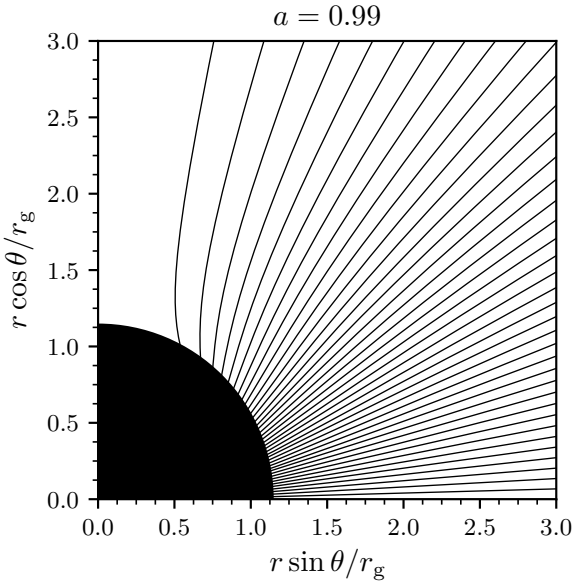


FIGURE 6.1: Magnetic field lines of a magnetic monopole in a Kerr spacetime with $a = 0.99$, described by Eq. (6.24).

$B^\varphi = 0$. Since we have already computed the electric and magnetic charges of the field F_η, it is straightforward to compute those associated with $*F_\eta$: the magnetic charge is $\overline{P} = \int *F_{\eta\theta\varphi}\,\mathrm{d}\theta\,\mathrm{d}\varphi/4\pi = P$, whereas the electric charge is $\overline{Q} = \int **F_{\eta\theta\varphi}\,\mathrm{d}\theta\,\mathrm{d}\varphi/4\pi = -\int F_{\eta\theta\varphi}\,\mathrm{d}\theta\,\mathrm{d}\varphi/4\pi = 0$. Again, thanks to the uniqueness theorem 6.1.1, we are assured that we have derived the only stationary, axisymmetric, vacuum solution of Maxwell's equations with no electric charge and a magnetic charge P. The magnetic field lines of the Kerr monopole are represented in Fig. 6.1. We used this solution to test the electromagnetic solver of `GRZeltron`, and as an initial condition in the simulations described in Chap. 12.

Dipoles and higher order

The "no-hair" theorem (as coined by J. A. Wheeler) states that a stationary black hole is only determined by its mass, angular momentum and electric charge (and its magnetic charge if such thing exists). The idea behind it is that any additional physical field is subject to stringent conditions, in order to reach equilibrium despite the tremendous gravity at the event horizon. Any perturbation will result in the black hole radiating away energy, so that the black hole rearranges itself until it reaches a stationary state (Frolov & Novikov, 1998). The only fields that remain are those subject to a conservation law and can be measured by a distant observer. So a "hair" actually means a configuration of the electromagnetic and gravitational fields which requires additional parameters to be described. Note that this theorem only applies to electromagnetic and gravitational fields: some unifying theories were able devise scalar fields that do not follow this property.

Still, black holes cannot support non-monopolar magnetic fields by their own. Imploding stars must lose their magnetic fields as they collapse through dipolar radiation. This does not contradict the presence of a magnetic moment carried by an electrically charged Kerr black hole, as the magnetic moment is in this case proportional to the electric charge and is not an independent parameter. However, black holes can live with non-monopolar magnetic fields supported by external currents, flowing outside of the event horizon. In that sense, black holes are very different from pulsars, which carry their own magnetic fields are can produce plasma by themselves.

6.1.3 Wald's solution

A special case of utmost importance consists in a black hole embedded in an external, asymptotically uniform magnetic field. This is an example of great astrophysical significance: one can expect that currents located in a remote accretion disk produce a large-scale magnetic field which is almost uniform at the scale of the central compact object. The vacuum solution was derived by Wald (1974) in the case of an external field aligned with the spin axis of the black hole. It was generalized to arbitrary inclinations by Bicak & Janis (1985).

In Sec. 6.1.2, we have studied the electromagnetic field generated by the timelike Killing vector $\boldsymbol{\eta}$, but the same can be done for $\boldsymbol{\psi}$. The covariant components of $\boldsymbol{\psi}$ are

$$\psi_\mu = (g_{0\varphi}, 0, 0, g_{\varphi\varphi}) = \left(-\frac{2ar_{\rm g}^2 r \sin^2\theta}{\Sigma}, 0, 0, \frac{A\sin^2\theta}{\Sigma} \right). \tag{6.25}$$

In the distant limit $r \gg r_{\rm g}$, these components reduce to $\psi_\mu \simeq \left(0, 0, 0, r^2\sin^2\theta\right)$. The associated magnetic field has components $B^r \simeq 2\cos\theta$ and $B^\theta \simeq -2\sin\theta/r$ (with $\sqrt{h} \simeq r^2\sin\theta$). Their form is more transparent in an orthonormal basis $\boldsymbol{e}_{\hat{i}} = \partial_i/\sqrt{h_{ii}}$. In the basis $\boldsymbol{e}_{\hat{i}}$, the components of the magnetic field are $B^{\hat{r}} = 2\cos\theta$ and $B^{\hat{\theta}} = \sqrt{h_{\theta\theta}}B^\theta = rB^\theta = -2\sin\theta$. Therefore, the 4-potential $A_{\psi_\mu} = (B_0/2)\psi_\mu$ generates an asymptotically uniform magnetic field of strength B_0. In order to make use of the uniqueness theorem 6.1.1, it is necessary to compute the electric and magnetic charges associated with this field. The electric charge Q_ψ enclosed in the sphere $\mathcal{S}_r$ is given by

$$4\pi Q_\psi = \int_{\mathcal{S}_r} {*F}_{\psi_{\theta\varphi}}\, \mathrm{d}\theta\, \mathrm{d}\varphi, \tag{6.26}$$

where ${*F}_{\psi_{\theta\varphi}}$ is determined just like in Eq. (6.14):

$${*F}_{\psi_{\theta\varphi}} = \sqrt{-g}F_\psi{}^{0r} = \sqrt{-g}\left(g^{00}g^{rr}F_{\psi_{0r}} + g^{0\varphi}g^{rr}F_{\psi_{\varphi r}}\right) \tag{6.27}$$

$$= \sin\theta\left(\frac{A}{\Sigma}\partial_r A_{\psi_t} + \frac{2r_{\rm g}^2 ra}{\Sigma}\partial_r A_{\psi_\varphi}\right). \tag{6.28}$$

At the lowest order in r, we have $A/\Sigma \simeq r^2$, $2ar_{\rm g}^2 r/\Sigma \simeq 2ar_{\rm g}^2/r$ and

$$\partial_r A_{\psi_0} \simeq B_0\frac{ar_{\rm g}^2\sin^2\theta}{r^2}, \quad \partial_r A_{\psi_\varphi} \simeq B_0 r\sin^2\theta. \tag{6.29}$$

This yields

$${*F}_{\psi_{\theta\varphi}} \simeq B_0 ar_{\rm g}^2\sin^3\theta + 2B_0 ar_{\rm g}^2\sin^3\theta = 3B_0 ar_{\rm g}^2\sin^3\theta. \tag{6.30}$$

All in all, the electric charge associated with A_ψ is

$$Q_\psi = \frac{1}{4\pi}\int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} 3B_0 ar_{\rm g}^2\sin^3\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi = 2aB_0 r_{\rm g}^2. \tag{6.31}$$

> Similarly, this integral is related to the *Komar angular momentum*, which is a global quantity of the spacetime and does not depend on the choice of $\mathcal{S}_r$ (Gourgoulhon, 2012). This is one of the most proper ways to define the angular momentum of spacetime.

The magnetic charge associated with A_ψ is given by $4\pi P_\psi = \int_{\mathcal{S}_r} F_{\psi\theta\varphi}\,\mathrm{d}\theta\,\mathrm{d}\varphi$. With $F_{\psi\theta\varphi} = (B_0/2)\partial_\theta\psi_\varphi = B_0 r^2 \sin\theta\cos\theta$, we see that the integral in θ vanishes. So the field A_ψ has no magnetic charge and an electric charge $2aB_0 r_\mathrm{g}^2$.

We are interested in constructing a stationary, axisymmetric electromagnetic field which tends to a uniform magnetic field of strength B_0 at infinity but carries neither an electric charge nor a magnetic charge. Thanks to the uniqueness theorem, we know that this solution is unique. To "cancel" the charge of A_ψ, we add a contribution from A_η, and define $A_\mathrm{W} = A_\psi + A_\eta$, with the constant factor Q to be adjusted. A_W tends to a uniform magnetic field at infinity, and its charge is $2aB_0 r_\mathrm{g}^2 + Q$ according to the Eq. (6.15) and (6.31). For the electric charge to vanish, we must have $Q = -2aB_0 r_\mathrm{g}^2$. All in all, using the expressions of A_η and A_ψ, the sought-after 4-potential is

$$A_{\mathrm{W}\mu} = \frac{B_0}{2}\left(\psi_\mu + 2ar_\mathrm{g}\eta_\mu\right) = \frac{B_0}{2}\left(g_{0\varphi} + 2ar_\mathrm{g}g_{00}, 0, 0, g_{\varphi\varphi} + 2ar_\mathrm{g}g_{0\varphi}\right). \tag{6.32}$$

This solution was derived in the BL coordinates, but remains the same in KS coordinates, since $A_r = A_\theta = 0$. The covariant components of the 4-potential are $A_\mathrm{W}{}^\mu = B_0\,(ar_\mathrm{g}, 0, 0, 1/2)$.

6.1.4 Properties of the Wald solution

Meissner effect

As we will see, a crucial quantity in order to determine how much energy is extracted from the black hole is the magnetic flux through one hemisphere of the event horizon Φ. Recalling Eq. (5.40), we have $\Phi = 2\pi A_{\mathrm{W}\varphi}\left(r_\mathrm{h}, \pi/2\right)$. Computing the coefficients of the metric at $r = r_\mathrm{h}$ and $\theta = \pi/2$, and using the identity $r_\mathrm{h}^2 - 2r_\mathrm{g}r_\mathrm{h} + a^2 r_\mathrm{g}^2 = 0$, we find that $g_{\varphi\varphi}\left(r_\mathrm{h}, \pi/2\right) = 4r_\mathrm{g}^2$ and $g_{t\varphi}\left(r_\mathrm{h}, \pi/2\right) = -2ar_\mathrm{g}^2/r_\mathrm{h}$. This yields $\Phi = \pi B_0 r_\mathrm{g}^2\left(4 - 4a^2 r_\mathrm{g}/r_\mathrm{h}\right)$. This can be reorganized in a perhaps more transparent form, using $r_\mathrm{h}^4 = (2r_\mathrm{g}r_\mathrm{h} - a^2 r_\mathrm{g}^2)^2 = 4r_\mathrm{h}^2 r_\mathrm{g}^2 + a^4 r_\mathrm{g}^4 - 4a^2 r_\mathrm{h} r_\mathrm{g}^3$:

$$4 - 4a^2\frac{r_\mathrm{g}}{r_\mathrm{h}} = \frac{4r_\mathrm{h}^2 - 4a^2 r_\mathrm{g}r_\mathrm{h}}{r_\mathrm{h}^2} = \frac{r_\mathrm{h}^4 - a^4 r_\mathrm{g}^4}{r_\mathrm{h}^2 r_\mathrm{g}^2}. \tag{6.33}$$

The flux through the upper hemisphere of the black hole reads

$$\Phi = \pi B_0 r_\mathrm{h}^2\left(1 - a^4\left(\frac{r_\mathrm{g}}{r_\mathrm{h}}\right)^4\right). \tag{6.34}$$

This result was first derived by King et al. (1975), in a rather involved manner. As the spin parameter a increases to its maximum value 1, the flux Φ tends to zero ($r_\mathrm{h} = r_\mathrm{g}$ for $a = 1$, $r_\mathrm{h} = 2r_\mathrm{g}$ for $a = 0$). For maximally rotating black holes, the magnetic field lines are totally expelled from the black hole, and they remain tangent to the event horizon (see Fig. 6.2). This is reminiscent of the Meissner effect: the expulsion of magnetic field lines from a superconductor that reaches a temperature below its critical temperature (see Collectif Jolidon 2021 for a pedagogical experimental realisation). This "Meissner" effect was invoked to dismiss the validity of the Blandford-Znajek mechanism (Bičák et al., 2007). As we will see later, this effect does not hold in the presence of plasma (Komissarov & McKinney, 2007).

Unscreened electric field

The spin of the black hole has an important consequence on the relative strengths of the electric and magnetic fields. In the Schwarzschild case, the electromagnetic field reduces to a vertical uniform magnetic field, since $A_{W\mu} = (B_0/2)\,(0, 0, 0, g_{\varphi\varphi}) = \left(0, 0, 0, B_0 r^2\sin^2\theta/2\right)$. In particular, the electric field $\boldsymbol{E}$ vanishes, so that $\boldsymbol{D} = -\boldsymbol{\beta}\times\boldsymbol{B}/\alpha$ and $\boldsymbol{D}\cdot\boldsymbol{B} = 0$. Although there might be an electric field as seen by FIDOs, depending on the choice of coordinates

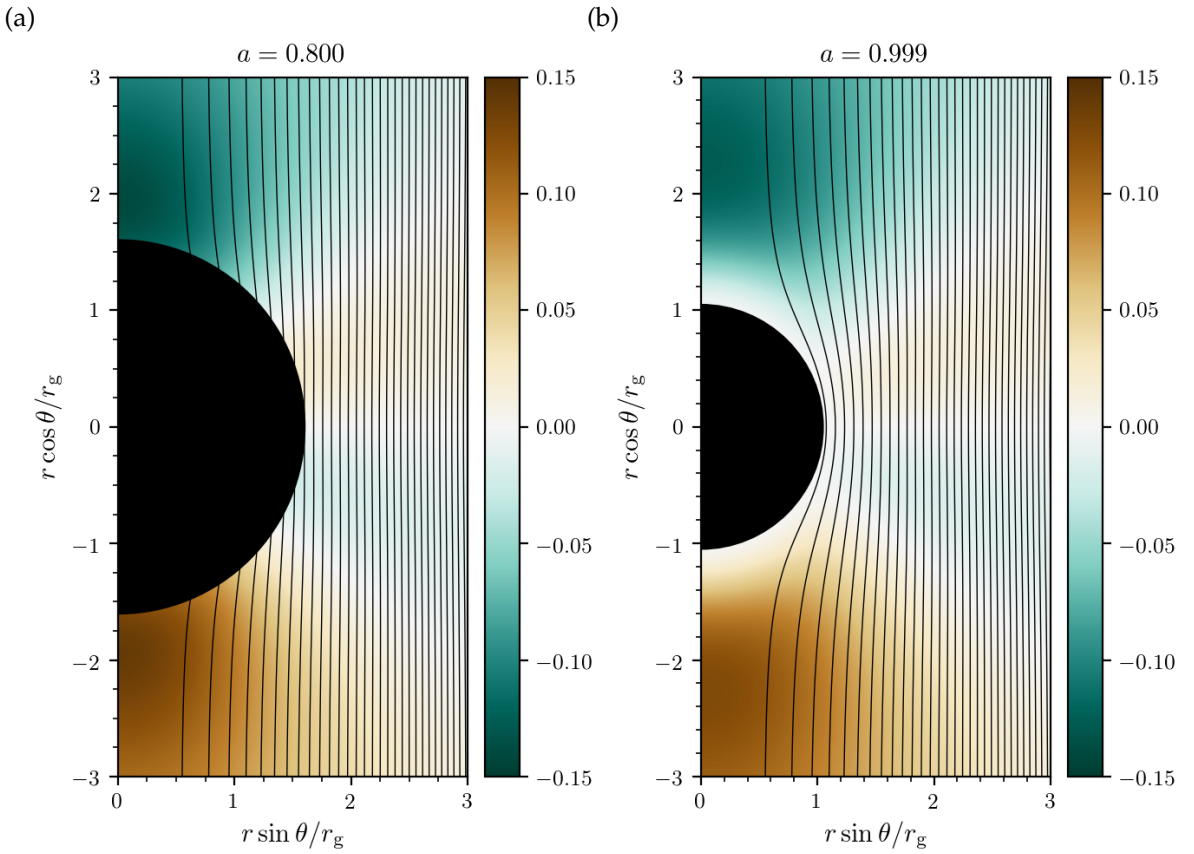


FIGURE 6.2: Parallel electric field $\boldsymbol{D}\cdot\boldsymbol{B}$ for a Kerr black hole with $a = 0.8$ (a) and $a = 0.999$ (b) in the Wald configuration, described by Eq. (6.32). The black solid lines are magnetic field lines.

and the shift vector, there can never be a parallel electric field. On the other hand, if $a \neq 0$, the components F_{0r}, $F_{0\theta}$, $F_{r\varphi}$ and $F_{\theta\varphi}$ do not vanish, so that the invariant

$$*F_{\mu\nu}F^{\mu\nu} = \frac{1}{2\sqrt{-g}}\left(F_{0r}F_{\theta\varphi} - F_{0\theta}F_{r\varphi}\right) = 4\,\boldsymbol{D}\cdot\boldsymbol{B} \tag{6.35}$$

does not vanish either (it is tedious, but possible, to check that there is no exact unfortunate compensation between $F_{0r}F_{\theta\varphi}$ and $F_{0\theta}F_{r\varphi}$; the components $F_{\mu\nu}$ for the Wald solution are given in BL coordinates by Bicak & Janis (1985) for example). The parallel electric field $\boldsymbol{D}\cdot\boldsymbol{B}$ is shown in Fig. 6.2. Thus, another major feature of the Wald solution is the occurrence of gravitationally induced electric fields which have a component parallel to the magnetic field, capable of accelerating particles and driving currents.

6.2 The importance of plasma in the force-free approximation

Let us reiterate the assumption of axisymmetry and time independence, that has already been essential in deriving the 3+1 field equations. Under such conditions, a vacuum magnetosphere is characterized by $\boldsymbol{\nabla}\times\boldsymbol{H} = \boldsymbol{0}$ and $\boldsymbol{\nabla}\times\boldsymbol{E} = \boldsymbol{0}$. Both these equations, supplemented with axisymmetry, imply $H_\varphi = E_\varphi = 0$. But we know from Eq. (5.55) that the Poynting flux is $\boldsymbol{\Pi} = (c/4\pi)\boldsymbol{E}\times\boldsymbol{H}$. If $H_\varphi = E_\varphi = 0$, then the poloidal component of $\boldsymbol{\Pi}$ vanishes, and there can be no extraction of energy from the black hole. We conclude that in order to extract energy from the black hole, the electromagnetic fields must be modified by electric charges

and currents. Not only are realistic astrophysical environments filled with plasma, but the presence of currents is compulsory in order to extract energy from the black hole. In this section, we will investigate how plasma impacts the properties of the magnetosphere.

6.2.1 The force-free approximation

An analytically convenient framework to alleviate the complexities of plasma behaviour is the force-free approximation, which is based on two opposite requirements. On the one hand, it is assumed that the density of the plasma is sufficiently high to screen all parallel electric fields. The requirement that the electric field should be entirely screened can be used to prove that the plasma density must be at least equal to the "Goldreich-Julian density". Anticipating a little bit, let us define as Ω the angular velocity of the magnetic field lines (a proper definition will be given later): the typical Goldreich-Julian charge density is defined as

$$n_{\rm GJ} = -\frac{\boldsymbol{\Omega}\cdot\boldsymbol{B}}{2\pi e c}, \tag{6.36}$$

with e the elementary charge. The ratio $\kappa = n/n_{\rm GJ}$ is called the "multiplicity" of the plasma. Electric screening requires charge separation of the plasma; however in the case of high multiplicities $\kappa \gg 1$, the plasma can be assumed quasi-neutral while the electric field remains screened.

On the other hand, the plasma is assumed to be so diffuse that its inertia is negligible. It is completely dominated by the Lorentz force, so that at equilibrium the volume Lorentz force must vanish. This translates into the requirement that the plasma magnetization σ should be very high, where σ is defined by

$$\sigma = \frac{\boldsymbol{B}^2}{8\pi n \gamma m c^2}, \tag{6.37}$$

In this equation, γ is the typical bulk Lorentz factor of the plasma and m the mass of the charged particles. A high magnetization implies that the plasma mass energy density is much lower than the electromagnetic energy density. The two force-free conditions

$$\begin{cases} \kappa \gg 1 & \text{(6.38a)} \\ \sigma \gg 1 & \text{(6.38b)} \end{cases}$$

must be met so the plasma can be considered force-free, but they have conflicting dependences with n. We need to know whether both conditions can be met, regardless of the plasma production mechanism for now. Let us evaluate an order of magnitude of σ for $n = \kappa n_{\rm GJ}$, in a mildly relativistic flow of electron/positron pairs of mass $m_{\rm e}$. Assuming assuming $\Omega \simeq c/r_{\rm g}$ and $\gamma \simeq 1$, we find

$$\sigma \simeq \frac{B_0^2}{8\pi\kappa n_{\rm GJ} m_{\rm e} c^2} \simeq \frac{B_0 e}{\kappa m_{\rm e} c \Omega} \simeq \frac{B_0 e r_{\rm g}}{\kappa m_{\rm e} c^2}. \tag{6.39}$$

Typical parameters for the supermassive black hole M87* are $B_0 \simeq 10\,\mathrm{G}$ and $r_{\rm g} \simeq 10^{15}\,\mathrm{cm}$, which yields $\sigma \simeq 10^{13}/\kappa$. Therefore, it is very possible to have both $\kappa \gg 1$ and $\sigma \gg 1$, if the magnetic field close to the magnetosphere is sufficiently strong.

6.2.2 Force-free electrodynamics

Assuming that the plasma conditions (6.38a) and (6.38b) are satisfied, we can write down the implications regarding the electromagnetic fields. The covariant force-free condition

reads (Blandford & Znajek, 1977; Komissarov, 2002b)

$$F_{\mu\nu}I^{\nu} = 0. \tag{6.40}$$

By very similar calculations as those from Sec. 5.3.1, one finds that the $3+1$ splitting of this equation is

$$\boldsymbol{E}\cdot\boldsymbol{J} = 0, \tag{6.41}$$

$$\rho\boldsymbol{E} + \frac{1}{c}\boldsymbol{J}\times\boldsymbol{B} = \boldsymbol{0}, \tag{6.42}$$

or equivalently

$$\boldsymbol{D}\cdot\boldsymbol{j} = 0, \tag{6.43}$$

$$\rho\boldsymbol{D} + \frac{1}{c}\boldsymbol{j}\times\boldsymbol{B} = \boldsymbol{0}. \tag{6.44}$$

The first equation states that there is no dissipation of electromagnetic energy: the electric field does not transfer energy to the plasma, and there is no particle acceleration. The second equation states that hydrodynamic forces are negligible, so that the volume Lorentz force vanishes. Taking the scalar product of this equation with $\boldsymbol{B}$ shows that $\boldsymbol{E}\cdot\boldsymbol{B} = \boldsymbol{D}\cdot\boldsymbol{B} = 0$. The electric field is screened, and has no component along the magnetic field. Recall that the screening occurs in the frame of a moving plasma element; in a laboratory frame there can be an inductive electric field due to the plasma motion across magnetic field lines. Eq. (6.44) also yields the poloidal FIDO-measured current density

$$\boldsymbol{j}_{\perp} = c\rho\frac{\boldsymbol{D}\times\boldsymbol{B}}{\boldsymbol{B}^2}. \tag{6.45}$$

This equation expresses that the perpendicular current is solely supported by the $\boldsymbol{E}\times\boldsymbol{B}$ drift of the charge-separated plasma. For these currents to be carried by a charge-separated plasma flowing at subluminal velocities, the current density must satisfy $\boldsymbol{j}_{\perp}^2 < \rho^2c^2$. Given that $\boldsymbol{D}\cdot\boldsymbol{B} = 0$, this leads to

$$\boldsymbol{B}^2 - \boldsymbol{D}^2 > 0. \tag{6.46}$$

> This condition can also be expressed with the $3+1$ form of the other Lorenz invariant associated with the Maxwell tensor: $F_{\mu\nu}F^{\mu\nu} > 0$.

In other words, the electromagnetic field must be magnetically dominated. If it were not, one could find a frame in which the electromagnetic field would be purely electric. The force-free approximation would break down: either the electric field would also vanish, or it would accelerate particles and induce currents *across* magnetic field lines (in the force-free framework, poloidal currents must flow along poloidal magnetic field lines).

Axisymmetry and time independence imply that $E_{\varphi} = 0$. As a result, both $\boldsymbol{E}$ and $\boldsymbol{J}\times\boldsymbol{B}$ are in the poloidal plane, and we know that $\boldsymbol{E}\cdot\boldsymbol{B} = \boldsymbol{D}\cdot\boldsymbol{B} = 0$. Therefore, one can define a 3-vector $\boldsymbol{\Omega} = \Omega\partial_{\varphi}$ such that

$$\boldsymbol{E} = -\frac{1}{c}\boldsymbol{\Omega}\times\boldsymbol{B}. \tag{6.47}$$

Ω is the angular velocity of the field lines, measured by an observer at infinity. This definition implies

$$\boldsymbol{D} = -\frac{1}{\alpha}\left(\frac{\boldsymbol{\Omega}}{c} + \boldsymbol{\beta}\right)\times\boldsymbol{B}. \tag{6.48}$$

Because $\omega = -c\beta^\varphi$ is the angular velocity of a BL FIDO, this last equation allows us to state that the angular velocity of field lines as measured by FIDOs is $\Omega - \omega$.

> To better grasp the definition (6.47) of Ω, let us look at the flat spacetime limit, for which $h_{\varphi\varphi} \simeq r^2 \sin^2\theta$. In the orthonormal spherical basis $\boldsymbol{e}_{\hat{i}}$, the electric field (6.47) reads
>
> $$\boldsymbol{E} = -\frac{\Omega}{c}\sqrt{h_{\varphi\varphi}}\boldsymbol{e}_{\hat{\varphi}} \times \boldsymbol{B} = -\frac{\Omega r \sin\theta}{c}\boldsymbol{e}_{\hat{\varphi}} \times \boldsymbol{B} = -\frac{1}{c}\left(\Omega \boldsymbol{e}_{\hat{z}} \times \boldsymbol{r}\right) \times \boldsymbol{B}. \tag{6.49}$$
>
> This means that $\boldsymbol{E}$ is determined by the flux freezing condition, with its drift velocity being the usual rigid rotation velocity $\Omega\boldsymbol{e}_{\hat{z}} \times \boldsymbol{r}$, at the angular velocity Ω. In other words, any observer moving at angular velocity Ω sees a vanishing electric field.

A priori, Ω is an arbitrary function of r and θ. However, in steady state, it follows from Eq. (5.37d) that $\boldsymbol{\nabla} \times (\boldsymbol{\Omega} \times \boldsymbol{B}) = \boldsymbol{0}$. With the vector calculus identity $\boldsymbol{\nabla} \times (\boldsymbol{\Omega} \times \boldsymbol{B}) = \boldsymbol{\Omega}(\boldsymbol{\nabla} \cdot \boldsymbol{B}) - \boldsymbol{B}(\boldsymbol{\nabla} \cdot \boldsymbol{\Omega}) + (\boldsymbol{B} \cdot \boldsymbol{\nabla})\boldsymbol{\Omega} - (\boldsymbol{\Omega} \cdot \boldsymbol{\nabla})\boldsymbol{B}$, Eq. (5.37c) and the independence of $\boldsymbol{B}$ and Ω with φ, there only remains

$$\boldsymbol{B} \cdot \boldsymbol{\nabla}\Omega = 0. \tag{6.50}$$

Therefore, Ω is constant along magnetic field lines. This fact is known as Ferraro's isorotation law: if it were violated, field lines would wind up and violate the assumption of axisymmetry. Ω does not necessarily correspond to any material angular velocity.

Finally, it is possible to derive another quantity which is invariant along magnetic field lines in the force-free approximation. Inserting the fact that $(\boldsymbol{J} \times \boldsymbol{B})_\varphi = 0$ into Eq. (5.37b) leads to $((\boldsymbol{\nabla} \times \boldsymbol{H}) \times \boldsymbol{B})_\varphi = (4\pi/c)(\boldsymbol{J} \times \boldsymbol{B})_\varphi = 0$, which means that

$$\boldsymbol{B} \cdot \boldsymbol{\nabla} H_\varphi = 0. \tag{6.51}$$

The quantity H_φ has a nice physical interpretation. Using Ampère's law in steady state $\boldsymbol{\nabla} \times \boldsymbol{H} = 4\pi\boldsymbol{J}/c$, we can compute from the Stokes' theorem the current I flowing through a circle $\mathcal{C}$ of radius r, subtending an angle θ with the spin axis:

$$I = \iint \boldsymbol{J} \cdot \mathrm{d}\boldsymbol{S} = \frac{c}{4\pi}\int_{\mathcal{C}} \boldsymbol{H} \cdot \mathrm{d}\boldsymbol{l} = \frac{c}{2}H_\varphi. \tag{6.52}$$

H_φ is proportional to the poloidal current through $\mathcal{C}$: the electric currents flow in magnetic field tubes. This high number of invariants along field lines is what makes the force-free framework quite tractable and powerful. In particular, it allows one to reduce the steady-state problem to a single partial derivative equation on A_φ (the Grad-Shafranov equation). We will see in Sec. 7.1 a special case of Grad-Shafranov equation in the absence of gravity.

Example

Let us come back to Eq. (6.36), now that all notions have been properly defined. The Goldreich-Julian density was introduced in the context of pulsar magnetospheres by Goldreich & Julian (1969). In flat spacetime, in the limit where the corotation speed is not relativistic, the electric charge density needed to screen the electric field is

$$\rho_{\rm GJ} = \frac{\boldsymbol{\nabla}\cdot\boldsymbol{E}}{4\pi} = -\boldsymbol{\nabla}\cdot\left(\frac{\Omega r\sin\theta}{4\pi c}\boldsymbol{e}_{\hat{\varphi}}\times\boldsymbol{B}\right) = -\frac{\Omega\boldsymbol{e}_{\hat{z}}\cdot\boldsymbol{B}}{2\pi c}. \tag{6.53}$$

Eq. (6.53) also informs us of which sign of charge is needed, depending on the magnetic field configuration. For instance, if $\boldsymbol{B}$ is a dipolar field, negative charges are needed near the rotation axis and positive charges at the equator.
This electric charge density is associated with the minimal plasma number density $n_{\rm GJ} = \rho_{\rm GJ}/e$ needed to screen the parallel electric field. Eq. (6.36) provides a satisfactory order of magnitude for the Goldreich-Julian plasma density. It is also possible to derive an expression in full general relativity, but Eq. (6.36) is sufficient for our purposes.

6.2.3 Gravitationally induced electric fields

We saw that currents must be driven in the magnetosphere to induce energy extraction. In order to drive such currents, for the vacuum to break down (via processes which will be described in Sec. 8), either the electromagnetic field should be purely electric ($F_{\mu\nu}F^{\mu\nu} < 0$), or the electric field should not be screened ($*F_{\mu\nu}F^{\mu\nu} \neq 0$). This allows us to understand the importance of the Wald solution: this is a vacuum, non-monopolar magnetic configuration of a Kerr black hole with a parallel electric field being produced naturally. This feature of the particular Wald's solution is actually more general, as proven by Komissarov (2004a).

Theorem 6.2.1. *There are no steady-state axisymmetric vacuum electromagnetic fields around a black hole, supported by external currents, which satisfy both* $\boldsymbol{D}\cdot\boldsymbol{B} = 0$ *and* $\boldsymbol{B}^2 - \boldsymbol{D}^2 > 0$ *everywhere along the magnetic field lines penetrating the ergosphere of the black hole.*

Let us assume that the first of these conditions is met ($\boldsymbol{D}\cdot\boldsymbol{B} = 0$) everywhere around the black hole, and let us prove that in that case the second one must fail. The magnetic field is generated by external distant sources, which are essentially not rotating around the black hole. Since Ω is constant along magnetic field lines, this means that $\boldsymbol{E} = \boldsymbol{0}$ along field lines penetrating the ergosphere, and $\boldsymbol{D} = -\boldsymbol{\beta}\times\boldsymbol{B}/\alpha$. A FIDO-measured electric field is generated even if the sources do not rotate. In BL coordinates, $\boldsymbol{\beta} = (0, 0, g_{0\varphi}/g_{\varphi\varphi})$ is azimuthal and $B_\varphi = H_\varphi/\alpha = 0$. This implies that $\boldsymbol{D}^2 = \boldsymbol{\beta}^2 B^2/\alpha^2$ since $\boldsymbol{\beta}$ and $\boldsymbol{B}$ are orthogonal. Then we have

$$\boldsymbol{B}^2 - \boldsymbol{D}^2 = \frac{\alpha^2 - \boldsymbol{\beta}^2}{\alpha^2}B^2. \tag{6.54}$$

The ergoregion, bounded by the limit of static observers, is defined by $g_{00} > 0$. According to Eq. (5.13), the ergoregion is characterized by $g_{00} = -\alpha^2 + \boldsymbol{\beta}^2 > 0$. Consequently, inside the ergosphere we have $\boldsymbol{B}^2 - \boldsymbol{D}^2 < 0$, completing the proof. The gravitationally induced vacuum electric field has to be stronger than the magnetic field, no matter the strength of the magnetic field.

Regardless of the structure of the magnetic field created by the distant sources, we showed that around a stationary vacuum black hole, either the electric field is unscreened and capable of driving currents in the magnetosphere, or the electromagnetic field is purely electric.

6.2.4 The role of plasma

Let us now relax the vacuum assumption, and work in the force-free framework. This discussion follows closely Komissarov (2004a) and Toma & Takahara (2014). There can be an azimuthal magnetic field generated by poloidal currents and field line rotation. From $\boldsymbol{D} = -(\boldsymbol{\Omega}/c + \boldsymbol{\beta}) \times \boldsymbol{B}/\alpha = -(\Omega - \omega)\,\boldsymbol{\psi} \times \boldsymbol{B}/c\alpha$, we get

$$\boldsymbol{D}^2 = h_{\varphi\varphi}\left(\frac{\Omega-\omega}{c\alpha}\right)^2\left(B_r B^r + B_\theta B^\theta\right) = h_{\varphi\varphi}\left(\frac{\Omega-\omega}{c\alpha}\right)^2\left(\boldsymbol{B}^2 - B_\varphi B^\varphi\right). \tag{6.55}$$

Noting that the light surface function f (Eq. (4.10)) can be rewritten as $f(\Omega, r, \theta) = g_{00} + 2g_{0\varphi}\Omega/c + g_{\varphi\varphi}\Omega^2/c^2 = -\alpha^2 + h_{\varphi\varphi}\left(\omega - \Omega\right)^2/c^2$, and that $B_\varphi = H_\varphi/\alpha$ in BL coordinates, we conclude that the Lorentz invariant $\boldsymbol{B}^2 - \boldsymbol{D}^2$ is given, in the force-free regime, by

$$\alpha^2\left(\boldsymbol{B}^2 - \boldsymbol{D}^2\right) = -\boldsymbol{B}^2 f\left(\Omega, r, \theta\right) + \left(\frac{\left(\Omega-\omega\right)H_\varphi}{c\alpha}\right)^2. \tag{6.56}$$

Let us start in a state with $\Omega = 0$ and $H_\varphi = 0$. As stated in Sec. 6.2.3, this means that $\boldsymbol{B}^2 - \boldsymbol{D}^2$ is negative everywhere in the ergosphere. Consequently, for field lines which penetrate the ergosphere, poloidal currents are generated (the term H_φ in Eq. (6.56)), giving rise to an azimuthal magnetic field and allowing $\boldsymbol{B}^2 - \boldsymbol{D}^2$ to be positive. The strong electric field, by a process which will be described later, injects charges which help screen the electric field. Yet, so far, $\boldsymbol{D}^2$ has not decreased.

When H_φ is non-zero, it is possible to show that this induces a flux of angular momentum along the magnetic field lines (Komissarov, 2004a), which inevitably results in rotation of the plasma and of the field lines. There cannot be a steady state with $H_\varphi \neq 0$ and $\Omega = 0$. As Ω grows, from Eq. (6.55), we see that eventually the strength of the electric field decreases. Also, the light surface function f departs from $-\alpha^2 + \boldsymbol{\beta}^2$ when Ω is non-zero. However, for angular velocities Ω which remain below ω_h, there always remains a zone in the ergosphere where $f\left(\Omega, r, \theta\right) > 0$ (see Fig. 4.2), which makes the generation of poloidal currents compulsory. All in all, we see that poloidal currents and azimuthal magnetic field are properties of the steady-state magnetosphere. Due to the presence of plasma, H_φ and $\boldsymbol{E}$ no longer vanish, so there can be energy extraction from the black hole. It should be pointed out that H_φ, which is constant along magnetic field lines, is only generated on field lines that penetrate the ergosphere. In other field lines, even though the initial electric field has an unscreened component, it is not strong enough to drive currents.

We also note that because the electric current distribution is invariant by reflection across the equatorial plane, according to Eq. (5.37b) $\boldsymbol{H}$ must be orthogonal to this plane, so that $H_\varphi = 0$ vanishes at the equatorial plane. There, the force-free condition $\boldsymbol{B}^2 - \boldsymbol{D}^2 > 0$ cannot hold. Any force-free magnetosphere must present a current sheet, allowing to close the electric current circuit, but where the force-free approximation which has made it appear in the first place breaks down. A sketch of the steady-state force-free magnetospheric structure is given in Fig. 6.3. Negative poloidal currents flow at high latitudes, producing a negative H_φ. The sign of H_φ allows extraction of energy. The electrical circuit is closed by the current sheet, which supports positive currents. These currents flow along the last field lines which cross the ergosphere.

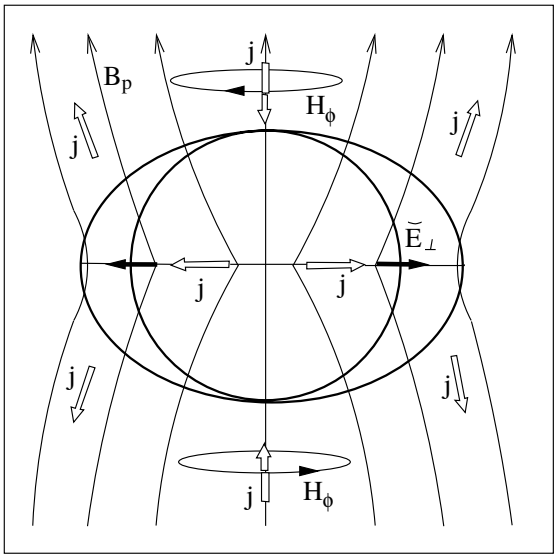


FIGURE 6.3: Electric current system in a steady-state magnetosphere. $\boldsymbol{j}$ denotes the electric current density, H_φ the toroidal magnetic field and $\boldsymbol{B}_\mathrm{p}$ the poloidal magnetic field. Figure taken from Komissarov (2004a).

Chapter 7

The Blandford-Znajek mechanism

Although we have derived several exact solutions of vacuum Maxwell's equations around a black hole, we have not been so prolific in the force-free case. As a matter of fact, only a few force-free solutions in a Kerr black hole magnetosphere exist. The Blandford-Znajek one is arguably the most consequential. Realizing the similarity between a Kerr black-hole and a neutron-star magnetosphere, they derived a solution in the limit of weak spin $|a| \ll 1$. Importantly, their solution allows the launch of an electromagnetically driven wind from the black hole, which extracts its energy and angular momentum. At last, the Penrose process got an astrophysically realistic implementation. In Sec. 7.1, we give a short overview of Michel's solution of a force-free pulsar magnetosphere. Aside from giving insight into the Blandford-Znajek solution, this flat spacetime solution will be used to derive some properties of the black-hole solution in Sec. 7.2.

7.1 Michel's solution of a force-free pulsar magnetosphere

Unlike black holes, neutron stars carry their own magnetic field, which is dipolar in a first approximation (although see Bilous et al. 2019). Besides, neutron stars are incredibly conductive, whereas black holes have no material surface and cannot anchor magnetic field lines. However, these two compact objects share a common feature: just like black holes embedded in an external uniform magnetic field, rotating neutron stars can develop strong unscreened electric fields in vacuum. Such electric fields are thought to lift charges from the neutron star surface. Then, pair creation can occur copiously by magnetic conversion or photon-photon annihilation (Harding & Lai, 2006). This motivates the use of the force-free framework to deal with pulsar magnetospheres.

Let us assume that the symmetry axis of the magnetosphere $\boldsymbol{e}_{\hat{z}}$ matches the spin axis of the neutron star. In this section, it turns out that cylindrical coordinates (R, φ, z) are more convenient. Even though we have already introduced some of the following concepts in Sec. 6.2.2, we will use the standard notations in this context. Because we assume axisymmetry, the magnetic field can be written in the orthonormal cylindrical basis $\left(\boldsymbol{e}_{\hat{R}}, \boldsymbol{e}_{\hat{\varphi}}, \boldsymbol{e}_{\hat{z}}\right)$ as

$$\boldsymbol{B} = \frac{1}{R}\boldsymbol{\nabla}\Psi \times \boldsymbol{e}_{\hat{\varphi}} + \frac{I}{R}\boldsymbol{e}_{\hat{\varphi}}, \tag{7.1}$$

where Ψ denotes the flux function. In our previous notations, Ψ is simply A_φ, and we know that it remains constant on magnetic field lines. On the other hand, using Ampère's law, we see that the poloidal current flowing through a surface lying on a circle of radius R, at height z, centered on the symmetry axis, is $cI/2$. As a result, I is nothing but H_φ, and is also considered as a function of Ψ. Because H_φ is constant along magnetic field lines, which can equivalently be labeled by A_φ, we will consider I as a function of Ψ rather than of R and z.

The statement that the magnetic field lines corotate at angular velocity Ω reads $\boldsymbol{E} = -\left(\Omega \boldsymbol{e}_{\hat{z}} \times \boldsymbol{r}\right) \times \boldsymbol{B}/c$. Ω is *a priori* dependent on Ψ, or on the latitude. For all field lines crossing the stellar surface, the angular velocity must be equal to the stellar angular velocity Ω_s

(excluding a differential rotation of the star). Inserting Eq. (7.1) into the force-free condition $\rho\boldsymbol{E} + \boldsymbol{J} \times \boldsymbol{B} = \boldsymbol{0}$, along with Maxwell's equations yields the pulsar equation (Michel, 1973a; Scharlemann & Wagoner, 1973; Endean, 1974):

$$\left(1 - \frac{R^2}{R_{\rm LC}^2}\right)\left(\frac{\partial^2\Psi}{\partial R^2} + \frac{\partial^2\Psi}{\partial z^2}\right) - \left(1 + \frac{R^2}{R_{\rm LC}^2}\right)\frac{1}{R}\frac{\partial\Psi}{\partial R} = -I(\Psi)\frac{\partial I}{\partial\Psi}, \tag{7.2}$$

where $R_{\rm LC} = c/\Omega_{\rm s}$ is the light cylinder of the neutron star, beyond which corotation at angular velocity $\Omega_{\rm s}$ is superluminal. This equation is singular at the light cylinder $R = R_{\rm LC}$. Some solutions can be found for $I = 0$, but they come with vanishing currents and energy flux. In order to extract energy from the neutron star, the solution must have electric currents which cross the light cylinder and connect the two domains $R < R_{\rm LC}$ and $R > R_{\rm LC}$. However, Michel (1973b) found another remarkable non-trivial solution, searching for a monopole configuration of the field lines, valid for $\theta \in [0, \pi/2[$.

Derivation of Michel's solution

Let us assume that $\Psi(\theta) = -\cos\theta = -z/\sqrt{R^2+z^2}$. After computing the partial derivatives, the terms involving Ψ in Eq. (7.2) read

$$\frac{\partial^2\Psi}{\partial R^2} + \frac{\partial^2\Psi}{\partial z^2} = \frac{z}{(R^2+z^2)^{3/2}} = \frac{1}{R}\frac{\partial\Psi}{\partial R}. \tag{7.3}$$

The pulsar equation reduces to

$$\left(\frac{R}{R_{\rm LC}}\right)^2 \frac{2z}{(R^2+z^2)^{3/2}} = I\frac{\partial I}{\partial\Psi}. \tag{7.4}$$

We know that I, hence the left-hand side too, only depends on Ψ. One can then realize that the left-hand side can be expressed as

$$I\frac{\partial I}{\partial\Psi} = \frac{2}{R_{\rm LC}^2}\cos\theta\sin^2\theta = -\frac{1}{R_{\rm LC}^2}2\Psi\left(1-\Psi^2\right). \tag{7.5}$$

From this, we obtain the current function: $I(\Psi) = \left(\Psi^2 - 1\right)/R_{\rm LC} \Rightarrow I(\theta) = -\sin^2\theta/R_{\rm LC}$ (the sign of I is enforced by $\boldsymbol{\Omega}\cdot\boldsymbol{B} > 0$ in the upper hemisphere).

From the functions $\Psi(\theta) = -\cos\theta$, $\Omega(\theta) = \Omega_{\rm s}$ and $I(\theta) = -\Omega_{\rm s}\sin^2\theta/c$, the electromagnetic field can be entirely reconstructed. Going back to an orthonormal spherical basis, their components read (with B_0 the magnetic strength at the stellar radius r_0)

$$\begin{cases} B^{\hat{r}} = B_0\dfrac{r_0^2}{r^2}; & \text{(7.6a)} \\ B^{\hat{\theta}} = 0; & \text{(7.6b)} \\ B^{\hat{\varphi}} = \dfrac{I}{R} = -B_0\dfrac{r_0}{R_{\rm LC}}\dfrac{r_0}{r}\sin\theta = \dfrac{r\sin\theta}{R_{\rm LC}}B^{\hat{r}}; & \text{(7.6c)} \end{cases}$$

whereas the electric field is simply $\boldsymbol{E} = B^{\hat{\varphi}}\boldsymbol{e}_{\hat{\theta}}$.

This solution is quite simple; the poloidal component is unchanged from an unrotating monopole. Magnetic field lines in the equatorial plane are spirals that wind up around the pulsar (see Fig. 7.1(a)). The toroidal field dominates over the radial field beyond the light cylinder. To satisfy $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$, this solution can only be valid in the upper hemisphere, the solution in the lower hemisphere being obtained via the transformation $\boldsymbol{B} \to -\boldsymbol{B}$, $\boldsymbol{E} \to -\boldsymbol{E}$. The reversal of the magnetic field across the equator implies the presence of an equatorial

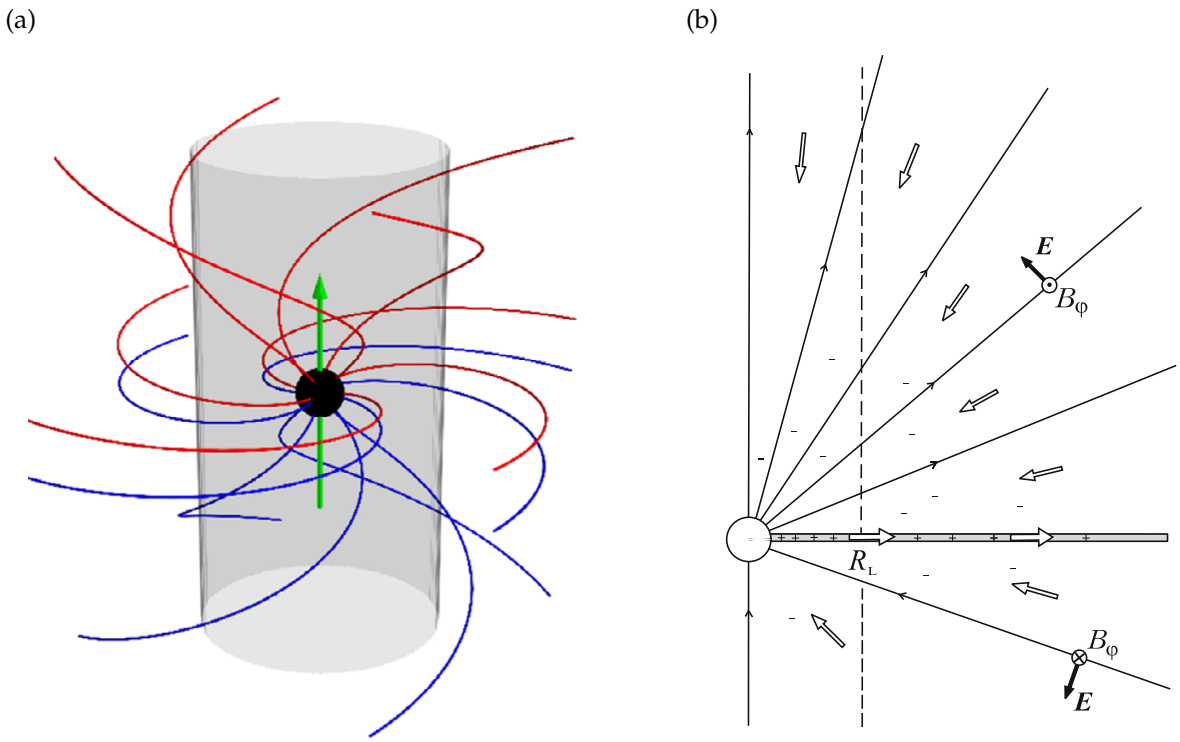


FIGURE 7.1: (a) Magnetic field lines in the analytical monopole solution. The green arrow is the rotation axis. Red lines represent outgoing field lines whereas blue lines are ingoing field lines. Figure taken from Cerutti & Beloborodov (2017). (b) Features of the split monopole solution. Magnetic field lines are solid black lines, whereas the electric current density is pictured by contour arrows. Figure taken from Beskin (2010).

current sheet. The current flowing in both hemispheres is negative, the closure of the circuit being allowed by this current sheet. The global structure of the magnetosphere is represented in Fig. 7.1(b). More importantly, this solution has a non-vanishing poloidal Poynting flux $\Pi^r = cE^{\hat{\theta}}B^{\hat{\varphi}}/4\pi$, with a total electromagnetic luminosity $L_s = (2/3)\Omega_s^2 B_0^2 r_0^4/c$ (Cerutti & Beloborodov, 2017). In this case, it is the conductive nature of the star surface that generates an electromotive force and allows this winding of field lines, hence this energy extraction.

Let us highlight an crucial point: although the plasma energy density is much smaller than the magnetic energy density, its presence is of utmost importance. The structure of the the fields in the far zone is strongly influenced by the currents the plasma carries. As the magnetic field is frozen in the plasma, which cannot rotate with superluminal velocity, field lines beyond the light cylinder are wrapped backwards with respect to the rotation of the star. It is the poloidal currents carried by the plasma which produce this toroidal field.

7.2 The Blandford-Znajek solution

7.2.1 A slowly rotating monopole

Michel's solution is an exact solution for any angular velocity of the neutron star. It can be easily extended to a Schwarzschild black hole; however, without rotation, no energy can be extracted. The goal of Blandford & Znajek (1977) was to find a similar energy-extracting solution in a Kerr black-hole force-free magnetosphere. Such a promising solution could not be found in the Kerr spacetime. Nonetheless, the similarities between black-hole and neutron star magnetospheres were striking, prompting Blandford and Znajek to investigate

a possible solution through a perturbative method. They expanded the field equations in powers of a. In the following, we will retain terms up to second order in a.

The most systematic approach would be to derive the general relativistic analog of the pulsar equation for Ψ (a Grad-Shafranov equation), then solve it for small a. We will take a more straightforward path, although less powerful mathematically. We know that Michel's solution must be valid in the flat spacetime limit, far from the black hole, with the caveat that the angular velocity of the field lines Ω is unknown. One needs to match this solution with what happens close to the black hole.

Znajek (1977) showed that on the event horizon, the toroidal field $H_\varphi\left(A_\varphi\right)$ must satisfy the following condition, valid for any spin parameter a:

$$H_\varphi\left(r_{\rm h},\theta\right)=\frac{\Omega-\omega_{\rm h}}{c}\,\frac{2r_{\rm h}r_{\rm g}\sin\theta}{\Sigma\left(r_{\rm h},\theta\right)}\,\partial_\theta A_\varphi. \tag{7.7}$$

A proof is provided in the Appendix 7.A. Although this condition looks like a boundary condition, it is actually a regularity condition, which can be understood in several ways. The only thing needed to derive it is the requirement that the fields, as measured by the FIDOs, should be regular at the event horizon, as they should since there is no true singularity at $r=r_{\rm h}$ (Komissarov, 2004a). Alternatively, in the force-free limit, the event horizon matches the critical surface of fast magnetosonic waves (Uzdensky, 2004, 2005), so the Znajek condition can also be seen as a regularity condition at the fast critical point. Znajek's condition could also be derived by requiring that the radial derivatives of A_φ should be finite at $r_{\rm h}$, and inject this condition in the general relativistic Grad-Shafranov equation.

Now, let us assume that the poloidal magnetic field is close to monopolar (we restrict ourselves to the domain $\theta\in[0,\pi/2[$): the flux function is $A_\varphi=B_0r_{\rm g}^2\left(1-\cos\theta+O\left(a^2\right)\right)$. The sign of a does not impact the poloidal magnetic field, so there are no even powers of a in the expansion. The associated poloidal magnetic field lines are straight lines emerging from the horizon, with magnetic field $B^r=\partial_\theta A_\varphi/\sqrt{h}=B_0\sin\theta r_{\rm g}^2\left(1+O\left(a^2\right)\right)/\sqrt{h}$. This implies that $\partial_\theta A_\varphi=B^r\sqrt{h}$ is constant on all magnetic field lines, to second order in a. Then, Eq. (7.7) reads

$$H_\varphi\left(r_{\rm h},\theta\right)=\frac{\Omega-\omega_{\rm h}}{c}\sin\theta\left(B^r\sqrt{h}\right)\left(r_{\rm h},\theta\right)\left(1+o\left(a^2\right)\right), \tag{7.8}$$

because $2r_{\rm h}r_{\rm g}/\Sigma\simeq 2r_{\rm g}/r_{\rm h}\simeq 1$ for $a\ll 1$ and $\left(\Omega-\omega_{\rm h}\right)\left(r_{\rm g}/c\right)$ is of the order a. Eq. (7.8) is *a priori* only valid on the event horizon. However, because Ω, $\sin\theta$ and $B^r\sqrt{h}$ are all constant on magnetic field lines, which all cross the event horizon in this case, this equation is actually valid for all r. We conclude that the current function H_φ is completely prescribed by the regularity condition on the horizon. On the other hand, there is no such prescription for the angular velocity of the field lines $\Omega\left(A_\varphi\right)$. The angular velocity is actually determined by the torque balance along the flux tube (see MacDonald & Thorne 1982 for a discussion).

> In the theory of pulsar magnetospheres, it is quite the opposite: the angular velocity of the field lines is fixed by the rotation of the star, whereas the current has to adjust to external conditions.

To determine the angular velocity Ω, we demand that this solution should match Michel's solution, which is valid in flat spacetime, at infinity. We found that $B_{\hat\varphi}=B^{\hat\varphi}=-r\sin\theta\left(\Omega/c\right)B^{\hat r}$. At infinity, we have $\sqrt{h}\simeq r^2\sin\theta$ and $B_{\hat\varphi}=H_{\hat\varphi}$ in an orthonormal basis. Finally, $H_{\hat\varphi}$ is related to the component in the natural coordinate basis H_φ by $H_\varphi=\sqrt{h_{\varphi\varphi}}H_{\hat\varphi}\simeq r\sin\theta H_{\hat\varphi}$. Thus, Michel's solution can be expressed in the coordinate basis as

$$H_\varphi\left(r\to+\infty,\theta\right)=-\frac{\Omega}{c}\sin\theta B^r\sqrt{h}\left(1+o\left(a^2\right)\right). \tag{7.9}$$

Matching Eq. (7.8), valid everywhere, and Eq. (7.9), valid at infinity, leads to

$$\Omega = \frac{1}{2}\omega_{\rm h} + o\left(a^2\frac{c}{r_{\rm g}}\right) \simeq \frac{1}{2}\omega_{\rm h}. \tag{7.10}$$

The angular velocity of the field lines is found to be uniform to second order in $ac/r_{\rm g}$. This is the condition under which the field lines can remain monopolar, and our solution can be considered global. Other solutions exist (Thoelecke et al., 2019), but those remain valid in limited regions only. We are searching for a solution which matches Michel's, since we are interested in a black hole which emits an energy outflow to infinity. Interestingly, Ω does not match the angular velocity of the black hole, as a symptom of the black hole's lack of a material surface. The toroidal field finally reads

$$H_\varphi = -\frac{1}{2}\frac{\omega_{\rm h}}{c}B_0 r_{\rm g}^2 \sin^2\theta. \tag{7.11}$$

7.2.2 Energy extraction

As we mentioned, in a force-free magnetosphere, extraction of energy requires poloidal currents. Let us compute the total electromagnetic energy flux. The Poynting vector reads (Eq. (5.55)):

$$\boldsymbol{\Pi} = \frac{c}{4\pi}\boldsymbol{E}\times\boldsymbol{H} = -\frac{1}{4\pi}\left(\boldsymbol{\Omega}\times\boldsymbol{B}\right)\times\boldsymbol{H} = \frac{1}{4\pi}\left(\left(\boldsymbol{H}\cdot\boldsymbol{B}\right)\boldsymbol{\Omega} - \left(\boldsymbol{H}\cdot\boldsymbol{\Omega}\right)\boldsymbol{B}\right), \tag{7.12}$$

so that the poloidal Poynting flux reads

$$\boldsymbol{\Pi}_{\rm p} = -\frac{1}{4\pi}\Omega H_\varphi \boldsymbol{B}_{\rm p}. \tag{7.13}$$

> Note that we need to evaluate the flux of energy-at-infinity, in order to explain how energy is transferred from the black hole to infinity. As a matter of fact, the Poynting flux measured by a BL FIDO $(c/4\pi)\boldsymbol{D}\times\boldsymbol{B}$ at the event horizon is directed toward the black hole (Komissarov, 2009; Beskin, 2010).

Let us evaluate the Poynting flux at the event horizon. That way, we only need to evaluate H_φ at $r = r_{\rm h}$, so we can afford to be agnostic about how the solution behaves at infinity and about the selected Ω. We can also assume the fields are reversed in the domain $\theta \in]\pi/2, \pi]$, so that we can integrate the Poynting vector from 0 to π, ignoring the current sheet. From Eq. (7.13) and (7.8), we get the total Blandford-Znajek luminosity

$$\begin{aligned}
L_{\rm BZ} &= \int_{\theta=0}^{\pi}\int_{\varphi=0}^{2\pi} \Pi^r\left(r_{\rm h},\theta\right)\sqrt{h}\,{\rm d}\theta\,{\rm d}\varphi && (7.14)\\
&= \frac{1}{2c}\int_0^\pi \Omega\left(\omega_{\rm h}-\Omega\right)\sin\theta\left(B^r\sqrt{h}\right)^2{\rm d}\theta && (7.15)\\
&= \frac{r_{\rm g}^4}{2c}B_0^2\Omega\left(\omega_{\rm h}-\Omega\right)\int_0^\pi \sin^3\theta\,{\rm d}\theta && (7.16)\\
&= \frac{2}{3}\frac{r_{\rm g}^2}{c^2}\Omega\left(\omega_{\rm h}-\Omega\right)\ B_0^2 c r_{\rm g}^2. && (7.17)
\end{aligned}$$

This equation is accurate to second order in a. This means that energy can be extracted from the black hole through this process if and only if $\Omega \in [0, \omega_{\rm h}]$. If the disk were to rotate faster than $\omega_{\rm h}$, or to counterrotate, it would deliver energy to the black hole. This is considered

unphysical, since energy would flow from infinity to the black hole. Besides, among all values of Ω which permit energy extraction, the one selected by the matching with Michel's solution allows for maximal $L_{\rm BZ}$. For $\Omega = \omega_{\rm h}/2$, the Blandford-Znajek power reads

$$L_{\rm BZ} = \frac{1}{6}\left(\frac{\omega_{\rm h} r_{\rm g}}{c}\right)^2 B_0^2 c r_{\rm g}^2 \tag{7.18}$$

Expansions of the Blandford-Znajek power to higher orders of a have been derived in the literature and confirmed numerically (McKinney & Gammie, 2004; Tchekhovskoy et al., 2010; Pan & Yu, 2015). The second-order expansion that we have derived here, however, has been shown to be surprisingly accurate up to values of a very close to 1 (Tchekhovskoy et al., 2010). It will prove sufficient to our purpose.

> Some other solutions were derived in the literature (Menon & Dermer, 2007), but those cannot extract energy from the black hole.

All in all, Blandford & Znajek (1977) found an astrophysical way to extract energy from a rotating black hole, provided it is embedded in a force-free magnetosphere. Let us give an order of magnitude for $L_{\rm BZ}$ for $a \ll 1$ ($\omega_{\rm h} \simeq ac/4r_{\rm g}$), working out the numbers:

$$L_{\rm BZ} = \frac{1}{96} a^2 B_0^2 c \left(\frac{GM}{c^2}\right)^2 = 6.8 \times 10^{40} a^2 \left(\frac{M}{10^9\,{\rm M}_\odot}\right)^2 \left(\frac{B_0}{10^2\,{\rm G}}\right)^2 {\rm erg\cdot s^{-1}}. \tag{7.19}$$

This energy output from supermassive black holes is large enough to account for the energy of the jets, and their radiation. We took a field strength consistent with observations in M87* (Neronov & Aharonian, 2007; Event Horizon Telescope Collaboration et al., 2019b).

All in all, the split monopole magnetic configuration is extremely similar to the simple picture we drew in Sec. 7.1. A toroidal field develops in both hemispheres with opposite signs. The discontinuity in H_φ is supported by an equatorial current sheet, where the magnetic field almost vanishes and the force-free approximation breaks down. The Poynting flux is directed outward in both hemispheres.

Interestingly, numerical studies show that this magnetic configuration is quite general. In particular, all field lines which cross the ergosphere end up crossing the event horizon (Komissarov & McKinney, 2007). At ergospheric scales, the magnetic configuration around the black hole is very close to a split monopole, regardless of the initial structure. This can be seen as a consequence of hoop stresses arising from the toroidal magnetic field. This is also consistent with the fact, proved in Sec. 7.2.4, that all field lines which cross the inner light surface cannot turn back, and must cross the event horizon. This is because field lines are frozen in the plasma, which can only flow toward the black hole within the inner light surface. We can conclude that the effect of the highly conductive plasma in the magnetosphere is to nullify the Meissner effect: magnetic field lines actually thread the black hole radially, and the magnetic flux through the upper hemisphere of the black hole does not vanish (although the magnetic flux through the entire event horizon is zero). The formula (7.18) that we derived is valid for a low spin, and for the specific case of a monopolar magnetic configuration. In general, one can write

$$L_{\rm BZ} = ck\left(\frac{\omega_{\rm h} r_{\rm g}}{c}\right)^2 \Phi^2, \tag{7.20}$$

where k is a numerical constant, which depends on the field line geometry, and $\Phi \simeq 2\pi r_{\rm g}^2 B_0$ is the magnetic flux through the upper hemisphere. This expression highlights the importance of the poloidal magnetic flux through the black hole (Tchekhovskoy et al., 2010; Parfrey

et al., 2015). The Blandford-Znajek solution with monopolar magnetic field lines is retrieved for $k = 1/(24\pi^2) \simeq 4.2 \times 10^{-3}$.

7.2.3 Interpretation

A reasonably useful analogy to grasp the extraction of energy and angular momentum from astrophysical rotators is the Faraday disk, which is an unipolar inductor. Consider a conductive metal disk of radius R, immersed in a magnetic field $\boldsymbol{B}_0$ along its symmetry axis, with electric wires connected to its axis and the rim of the disk[1] (see a sketch in Fig. 7.2). Let us suppose at first that the electrical circuit is open. As the conducting rotator is set in motion to the angular velocity $\boldsymbol{\Omega}$, its charge carriers (usually electrons in a metal, but in a plasma positive charges can participate too) are forced into rotation by collisions with the lattice, so they experience an unbalanced Lorentz force. Alternatively, in the rotating frame, one can consider that they experience an inductive electric field $\boldsymbol{E}' = \boldsymbol{V} \times \boldsymbol{B}_0/c$, with $\boldsymbol{V} = \boldsymbol{\Omega} \times \boldsymbol{r}$ (primed quantities denote quantities measured in the rotating frame of the conductor, while unprimed quantities are measured in the laboratory frame). This is a consequence of the transformation of an electric field under a change of frames (Garing, 1999). This electromotive force induces a charge redistribution, with positive charges concentrating toward the rim, and negative charges converging at the axis. This motion stops when the electric field induced by the charge separation exactly compensates the Lorentz force, which occurs when the total electric field in the rotating frame vanishes: $\boldsymbol{E}' = \boldsymbol{E} + \boldsymbol{V} \times \boldsymbol{B}_0/c = \boldsymbol{0}$. An electric field $\boldsymbol{E}$ arises in the laboratory frame, so a potential difference U develops between the axis and the rim.

> Note that despite the its denomination as a Faraday disk, the principle of a unipolar inductor is not Faraday's law (an electromotive force induced by variations of magnetic flux through a closed loop). The "inducing" is done by Lorentz forces in this case. Useful discussions can be found in Michel & Li (1999) and Beskin (2010).

Now let us short out the circuit by electrically connecting the axis and the rim (so it becomes "unipolar"). Charges are free to flow from one end to the other. Forcing the rotation keeps inducing an electromotive force, which keeps driving currents. These currents are dissipated elsewhere in the system, thus extracting rotational energy from the disk. Let us give an order of magnitude for the dissipated energy: the electromotive force $e = -U$ is obtained by integration the electromotive field along any line between the axis and the rim (Garing, 1999)

$$e = \frac{1}{c} \int_{\text{axis}}^{\text{rim}} \boldsymbol{V} \times \boldsymbol{B} \cdot \mathrm{d}\boldsymbol{l} = -\frac{1}{2c} \Omega R^2 B_0. \tag{7.21}$$

The actual numerical prefactor is unimportant. The total dissipated power is $\mathcal{P} = e^2/\mathcal{R}$, with $\mathcal{R}$ the resistance of the system. For simplicity, let us assume that $\mathcal{R} = 4\pi/c$ is the impedance of free space; this is the most natural choice in the absence of any physical models, and this is the resistance which is ascribed to the event horizon in the membrane paradigm (Thorne et al., 1986). Then the dissipated power is

$$\mathcal{P} = \frac{c}{4\pi} e^2 \propto \frac{\Omega^2 B_0^2 R^4}{c}. \tag{7.22}$$

Up to a numerical factor, this is exactly the output electromagnetic power from a force-free neutron-star magnetosphere, or the Blandford-Znajek power (7.13). In this analogy, the conducting disk is replaced by the rotating, magnetized compact object. The highly conductive plasma serves to close the electrical circuit. The disk is submitted to a braking

[1] An example of practical implementation of a Faraday disk is the Barlow wheel.

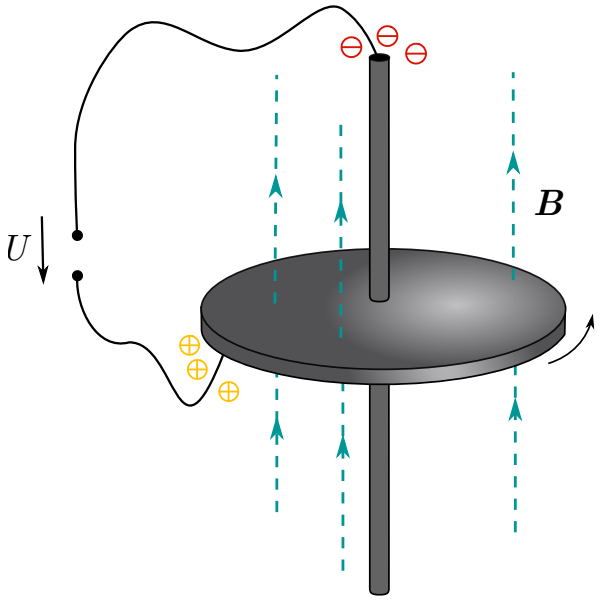


FIGURE 7.2: Sketch of a unipolar inductor. A conducting disk is threaded by magnetic field lines (blue dashed lines). As it is spun, the free charges in the disk redistribute (yellow and red circles), so that a potential drop U develops between the edge of the disk and its axis.

torque $\mathcal{M} \propto -\Omega B_0^2 R^4/c$, due to the Lorentz force exerted on the currents flowing on its surface, and must provide the power $\mathcal{P} = -\mathcal{M}\Omega$.

There are additional subtleties in the case of black holes, though. Black holes lack a material surface to anchor the magnetic field lines. The electromotive force cannot be interpreted as being due to charge separation in a rotating material conductor. Instead, spacetime itself is rotating. What drives the currents is the electric field generated gravitationally, rather than rotationally (see Eq. (5.41a)), by the black hole immersed in a magnetic field. The presence of the ergoregion is key to this process, a zone where it is impossible for the magnetic field to be stronger than the electric field. To develop currents and an electromagnetic energy outflow, what matters is whether field lines cross the ergosphere or not. The ergosphere would be the actual unipolar inductor, in this picture. Field lines which do not cross the ergosphere are not set into rotation at all.

Besides, the dependence in Ω in Eq. (7.17) is peculiar, and different from that of pulsars. This equation states that the extraction of energy is possible only if $0 < \Omega < \omega_\mathrm{h}$. The black hole is submitted to a "torque" $\mathcal{M} \propto -(\omega_\mathrm{h} - \Omega)\, B_0^2 r_\mathrm{g}^2/c$, with a different dependence on Ω than in the neutron star/Faraday disk case. The difference is that spacetime itself is already rotating at the angular velocity $\simeq \omega_\mathrm{h}$. If $\Omega \simeq 0$, no electromotive force is induced "seen from infinity", although a FIDO could measure one. On the other hand, if $\Omega \simeq \omega_\mathrm{h}$, no poloidal currents need develop because the field lines are at rest with respect to spacetime. Eq. (7.17) takes both effects into account, so that the energy transfer is maximal for an intermediate angular velocity, with field lines lagging behind the black hole. This bizarre behaviour is another side effect of spacetime dragging, which makes classical analogies break down somehow. MacDonald & Thorne (1982) give another classical interpretation in terms of an electrical circuit.

7.2.4 The importance of light surfaces

Before moving on to the next chapter, we will reinterpret the energy extraction condition $0 < \Omega < \omega_\mathrm{h}$ in a way that highlights the essential role of light surfaces (introduced in Sec. 4.1.5) in the structure of the magnetosphere. A magnetized plasma usually supports three kind of

waves: slow magnetosonic, Alfvén and fast magnetosonic waves. In the force-free limit, the velocity of slow waves goes to zero because the plasma has vanishing inertia, whereas fast waves travel at the speed of light. As mentioned earlier, the critical surface of fast waves is therefore located at the event horizon. The velocity of Alfvén waves travelling along the unit vector $\boldsymbol{n}$ in the force-free limit, either parallel or antiparallel to $\boldsymbol{n}$, is (Komissarov, 2002b)

$$v_{\pm} = \frac{c}{\boldsymbol{B}^2}\left(\boldsymbol{D}\times\boldsymbol{B} \pm \sqrt{\boldsymbol{B}^2 - \boldsymbol{D}^2}\boldsymbol{B}\right)\cdot\boldsymbol{n}. \tag{7.23}$$

We have already proved in Sec. 6.2.4 that $\alpha^2\left(\boldsymbol{B}^2 - \boldsymbol{D}^2\right) = -\boldsymbol{B}^2 f\left(\Omega, r, \theta\right) + H_{\varphi}^2\left(\omega - \Omega\right)/c^2\alpha^2$. In addition to that, from $\boldsymbol{D} = \left(\omega - \Omega\right)\left(\boldsymbol{\psi}\times\boldsymbol{B}\right)/c\alpha$, one obtains the expression of the cross product $\boldsymbol{D}\times\boldsymbol{B} = \left(\Omega - \omega\right)\left(\boldsymbol{B}^2\boldsymbol{\psi} - B_{\varphi}\boldsymbol{B}\right)/c\alpha$, with $B_{\varphi} = H_{\varphi}/\alpha$ in BL coordinates. Restricting ourselves to the case of a direction of propagation $\boldsymbol{n}$ in the poloidal plane, we get the expression of $v_{\pm}$:

$$v_{\pm} = \frac{c}{\boldsymbol{B}^2}\left(\boldsymbol{n}\cdot\boldsymbol{B}\right)\eta_{\pm}, \quad \eta_{\pm} = \frac{\omega - \Omega}{c}\frac{H_{\varphi}}{\alpha^2} \pm \sqrt{\left(\frac{\left(\omega - \Omega\right)H_{\varphi}}{c\alpha^2}\right)^2 - \frac{\boldsymbol{B}^2}{\alpha^2}f\left(\Omega, r, \theta\right)}. \tag{7.24}$$

Now, let us assume a fixed outgoing direction for the poloidal field $\boldsymbol{B}$: the condition $\boldsymbol{n}\cdot\boldsymbol{B} \neq 0$ tells us that Alfvén waves propagate along poloidal field lines. A critical Alfvén surface is defined as the locus where either v_+ or v_- changes sign. From Eq. (7.24), recalling that $H_{\varphi} < 0$, one sees that v_+ becomes negative below the inner light surface, because in this region $f\left(\Omega, r, \theta\right) > 0$ and $\Omega < \omega$ (see Fig. 4.2). Similarly, v_- becomes positive beyond the outer light surface, because in this region $f\left(\Omega, r, \theta\right) > 0$ and $\Omega > \omega$. This means that in the force-free limit, the light surfaces exactly match the Alfvén surfaces. Beyond the outer critical surface, all Alfvén waves must propagate outward, whereas all Alfvén waves emitted below the inner light surface propagate toward the black hole.

This is a stark difference with pulsar magnetospheres, which only have one critical surface (the outer light cylinder). The Blandford-Znajek solution in the Kerr metric possesses two critical surfaces, so it must describe both an outgoing and an ingoing wind. An important consequence of this fact is that there must be a plasma injection mechanism, overlooked in this section, which replenishes the magnetosphere with fresh plasma so it can remain force-free. This injection must take place between the two light surfaces, whereas in neutron

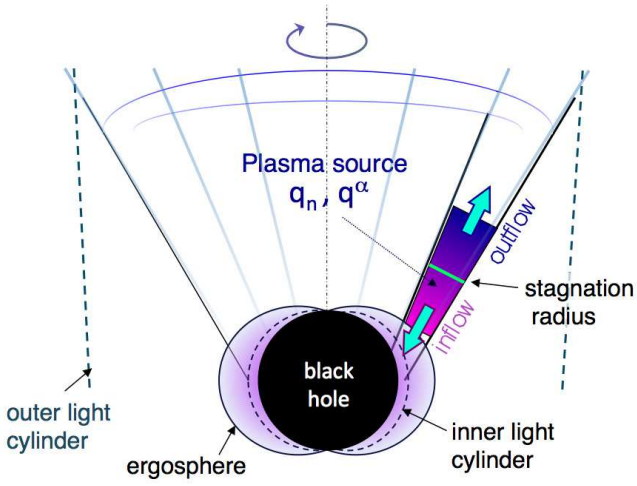


FIGURE 7.3: Sketch of the flow structure along a radial magnetic field line. A plasma source must lie between the two light surfaces. Figure taken from Globus & Levinson (2013).

star magnetospheres it could be injected from the stellar surface. This situation is sketched in Fig. 7.3. More will be said about this plasma supply mechanism in Sec. 12.1.

Furthermore, causality requires that the driving force responsible for launching the winds should be able to communicate with both the ingoing and outgoing winds. Punsly (2001) and Komissarov (2002a) showed that fast waves alone are not sufficient to build the poloidal current structure, and that the communication takes place by the means of both fast waves and Alfvén waves. We have already identified the ergosphere as the driving force of the energy extraction mechanism: poloidal currents must develop only on field lines which cross the ergosphere. Following on this idea, Takahashi et al. (1990) proved, in the MHD framework, the following theorem (see also Levinson 2006 for a discussion):

Theorem 7.2.1. *There can be rotational energy extraction from the black hole only if the inner Alfvén surface is located in the ergoregion.*

This is the case in Fig. 4.3. However, it turns out that if $\Omega < 0$, the inner light surface lies outside the ergosphere (see Fig. 7.4(a)). Rrecall that for $\Omega = 0$, the inner light surface is the ergosphere. On the other hand, if $\Omega > \omega_{\rm h}$, then the inner and outer light surfaces merge (see Fig. 7.4(b)). From Eq. (7.24), we see that $v_{\pm}$ is negative everywhere: Ω is greater than the velocity of the FIDO ω (see Fig. 4.2 for $\Omega > \omega_{\rm h}$) and H_{φ} is positive. Consequently, there can be no outflow of energy. So actually, in the force-free limit, the theorem 7.2.1 is equivalent to the energy extraction condition that was already derived: $0 < \Omega < \omega_{\rm h}$. This allows us to interpret this condition as the necessary condition for there to be a causal connection between the ergosphere and the ingoing and outgoing outflows (the only exception

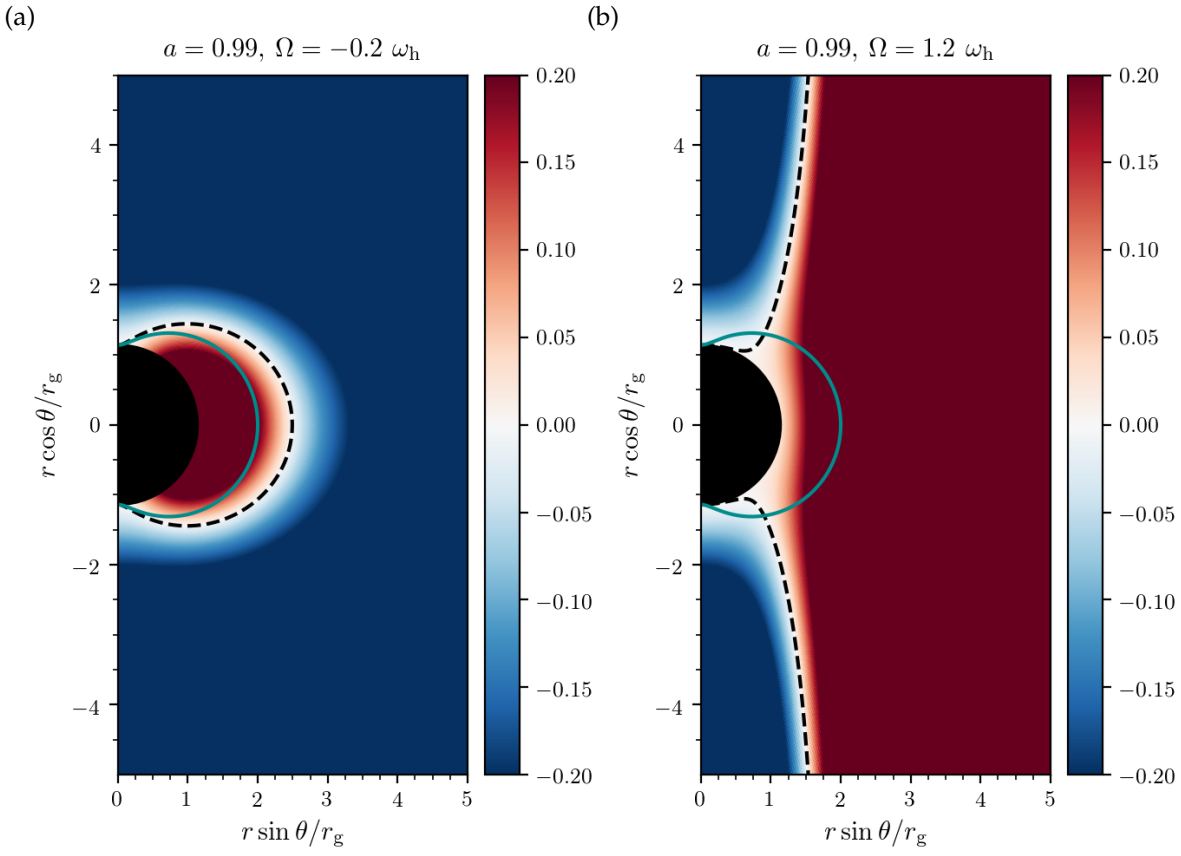


FIGURE 7.4: Spatial map of the light surface function f (Eq. (4.10)) for $a = 0.99$ and fixed $\Omega = -0.2\,\omega_{\rm h}$ (a) and $\Omega = 1.2\,\omega_{\rm h}$ (b).

is the polar direction $\theta = 0$, where the inner light surface intersects the ergosphere, but the Poynting flux vanishes at $\theta = 0$).

> Incidentally, we proved that if energy is extracted from the black hole, its event horizon is causally disconnected from the outgoing wind. This renders any interpretation of the Blandford-Znajek mechanism as a mechanical exerted on the event horizon improper, although this would be a valid interpretation of the spindown of neutron stars. Actually, Lasota et al. (2014) proved that tapping the rotational energy requires that the black hole should absorb a negative energy flux, making the Blandford-Znajek process the electromagnetic version of the Penrose process.

7.A Appendix: Proof of the Znajek regularity condition

Let us derive Eq. (7.7) in the 3+1 formalism, as the original proof from Znajek (1977) required an advanced formalism. The goal will be to express the r and φ components of $\boldsymbol{B}$ along a magnetic field line, using the quantities H_φ and Ω which we proved to be invariant on a single field line. Then we will ask that there should be no singularity at $r = r_\mathrm{h}$. For this procedure to be valid, the coordinate system that we use must not itself be singular at the event horizon; therefore we will carry out this calculation in the KS coordinate system (see the Appendix 4.A). Recall that in this system, the shift vector is purely radial: $\boldsymbol{\beta} = \beta^r \partial_r$, with $\beta^r = z/(1+z)$, $z = 2r_\mathrm{g} r/\Sigma$. In the following of this section, we will keep $r_\mathrm{g} = 1$ and $c = 1$ for the sake of conciseness, and we will denote the KS coordinates (t, r, θ, φ), without the hats.

Inserting Eq. (6.48) in Eq. (5.41b), we obtain

$$\boldsymbol{H} = \alpha \boldsymbol{B} + \left((\boldsymbol{\beta}\cdot\boldsymbol{B})\,\boldsymbol{\Omega} - \boldsymbol{\beta}^2 \boldsymbol{B} - (\boldsymbol{\beta}\cdot\boldsymbol{\Omega})\,\boldsymbol{B} + (\boldsymbol{\beta}\cdot\boldsymbol{B})\,\boldsymbol{\beta} \right)/\alpha. \tag{7.25}$$

This allows us to get expressions for H^φ and H^r, using $\boldsymbol{\beta}\cdot\boldsymbol{\Omega} = h_{ij}\beta^i\Omega^j = h_{r\varphi}\beta^r\Omega = -a\,z\,\Omega\sin^2\theta$, $\boldsymbol{\beta}^2 = h_{rr}\beta^{r2} = z^2/(1+z)$ and $\boldsymbol{\beta}\cdot\boldsymbol{B} = h_{rr}\beta^r B^r + h_{r\varphi}\beta^r B^\varphi = z\left(B^r - a\sin^2\theta B^\varphi\right)$:

$$H^\varphi = \alpha B^\varphi + \frac{1}{\alpha}\Omega\, z\, B^r - \frac{1}{\alpha}\frac{z^2}{1+z} B^\varphi; \tag{7.26}$$

$$H^r = \alpha B^r + \frac{1}{\alpha} a\,\Omega\, z \sin^2\theta B^r - \frac{1}{\alpha}\frac{z^2}{1+z} a\sin^2\theta B^\varphi. \tag{7.27}$$

Using $\alpha = 1/\sqrt{1+z}$ in KS coordinates, the covariant component H_φ can then be reconstructed:

$$H_\varphi = h_{\varphi\varphi}H^\varphi + h_{\varphi r}H^r \tag{7.28}$$

$$= \frac{A\sin^2\theta}{\Sigma}\left(\frac{B^\varphi}{\sqrt{1+z}} + \Omega z\sqrt{1+z}B^r - B^\varphi\frac{z^2}{\sqrt{1+z}}\right)$$

$$- a\sin^2\theta\,(1+z)\left(\frac{B^r}{\sqrt{1+z}} + a\,z\sqrt{1+z}\sin^2\theta B^r\Omega - a\sqrt{1+z}\sin^2\theta\frac{z^2}{1+z}B^\varphi\right) \tag{7.29}$$

$$= B^\varphi\sqrt{1+z}\sin^2\theta\left(\frac{A}{\Sigma}(1-z) + a^2\sin^2\theta z^2\right)$$

$$+ B^r\sqrt{1+z}\sin^2\theta\left(\Omega z\left(\frac{A}{\Sigma} - a^2\sin^2\theta\,(1+z)\right) - a\right) \tag{7.30}$$

This expression can be simplified, by making use of the identity

$$A = \left(r^2 + a^2\right)^2 - a^2\Delta\sin^2\theta = \Sigma\left(r^2 + a^2\right) + 2a^2 r\sin^2\theta. \tag{7.31}$$

This allows us to rewrite the coefficient of B^φ in Eq. (7.30) on the right-hand side as

$$\frac{A}{\Sigma}(1 - z\,) = \frac{1}{\Sigma^2}\left(A\left(\Sigma - 2r\right) + 4a^2r^2\sin^2\theta\right) \tag{7.32}$$

$$= \frac{1}{\Sigma}\left(r^4 + a^2r^2\cos^2\theta - 2r^3 + a^2r^2 + a^4\cos^2\theta - 2a^2r\cos^2\theta\right) \tag{7.33}$$

$$= \frac{1}{\Sigma}\left(r^2 - 2r + a^2\right)\left(r^2 + a^2\cos^2\theta\right) = \Delta. \tag{7.34}$$

Using the identity Eq. (7.31), one finds similarly the coefficient of B^r in Eq. (7.30):

$$\Omega z\left(\frac{A}{\Sigma} - a^2\sin^2\theta\,(1 + z)\right) - a = \Omega\frac{2r}{\Sigma^2}\left(A - \left(\Sigma + 2r\right)a^2\sin^2\theta\right) - a \tag{7.35}$$

$$= \Omega\frac{2r}{\Sigma^2}\left(\Sigma\left(r^2 + a^2\right) - \Sigma a^2\sin^2\theta\right) - a = 2r\Omega - a. \tag{7.36}$$

Now we can write the relation between B^r and B^φ on a given field line, with fixed Ω and H_φ:

$$B^\varphi\Delta\sin^2\theta = \alpha H_\varphi - B^r\sin^2\theta\,(2r\Omega - a)\,. \tag{7.37}$$

Although the coefficients of the metric are not singular at the event horizon, $\Delta(r)$ vanishes for $r = r_\mathrm{h}$. For B^φ to remain finite, the right hand-side of Eq. (7.37) must vanish as well, leading to

$$H_\varphi\left(r_\mathrm{h}, \theta\right) = \frac{\sin^2\theta}{\alpha\left(r_\mathrm{h}, \theta\right)}\left(2r_\mathrm{h}\Omega - a\right)B^r\left(r_\mathrm{h}, \theta\right). \tag{7.38}$$

This expression can still be made more convenient. We recall that the φ component of a 1-form is invariant under a change from BL to KS coordinates; however Eq. (7.38) is only true in KS coordinates. This means that H_φ and A_φ have the same expression in BL and KS coordinates (r and θ are also left unchanged). B^r can be expressed as $B^r = e^{rjk}\partial_j A_k = \partial_\theta A_\varphi/\sqrt{h}$. We use Eq. (5.19) to express the determinant h as a function of α. In KS coordinates, the determinant of the metric has the same expression $-g = \Sigma^2\sin^2\theta$ than in BL coordinates. Using the expression of the angular velocity of the black hole Eq. (4.14), and putting back the dimensional factors, we finally arrive at the Znajek condition (Znajek, 1977):

$$H_\varphi\left(r_\mathrm{h}, \theta\right) = \frac{\Omega - \omega_\mathrm{h}}{c}\,\frac{2r_\mathrm{h}r_\mathrm{g}\sin\theta}{\Sigma\left(r_\mathrm{h}, \theta\right)}\,\partial_\theta A_\varphi. \tag{7.39}$$

This condition is also valid in BL coordinates.

Chapter 8

Radiative processes

In the previous sections, we have deliberately ignored the important question of the origin of the plasma in a black-hole magnetosphere. In pulsars, free charges are supplied by the stellar surface, but in the present case there is no such obvious plasma source. Blandford & Znajek (1977), in their seminal paper, suggested that photon-photon annihilation could provide the right number of charges so as to reach the Goldreich-Julian density. Note that the detection of the $511\,\mathrm{keV}$ electron/positron annihilation line is evidence for the production of positrons in the Galaxy. In this section, we will describe several radiative processes which will be crucial to our work. Inverse Compton scattering is responsible for the production of high-energy photons, which can then produce electron/positron (hereafter, $e^{\pm}$) pairs. Synchrotron radiation probably produces most of the emission that we detect from AGN in radio and millimeter wavelength.

8.1 Inverse Compton scattering

8.1.1 Kinematics

Compton scattering describes the following free-free process between an electron (or a positron) and a photon:

$$\gamma + e^{-} \to \gamma + e^{-}. \tag{8.1}$$

Inverse Compton scattering is predominant in astrophysical environments, when both high-energy electrons/positrons and intense external radiation fields are present and share the same volume. In inverse Compton scattering, high-energy electrons scatter off low-energy photons, boosting them to higher energies. Instead of the photon losing energy, as would be the case in direct Compton scattering, here the electron loses energy. Let us denote, in the observer frame $\mathcal{R}$, by γ_0 (resp. γ_1) the Lorentz factor of the electron before (resp. after) the scattering, and similarly the energies ϵ_0 and ϵ_1 of the photon before and after the scattering. Primed quantities will be used to denote quantities defined in the electron rest frame $\mathcal{R}'$. For simplicity, let us assume that the electron moves along the z axis, in a frame with spherical coordinates (r, θ, φ) defined with respect to this axis. We also suppose the electron collides the photon with an angle θ_0 in the observer frame. The convention we use for the angles is shown in Fig. 8.1.

The photon energy in $\mathcal{R}'$ is $\epsilon'_0 = \gamma_0 \epsilon_0 \left(1 - \beta \cos\theta_0\right)$, with $\beta_0^2 = \left(\gamma_0^2 - 1\right)/\gamma_0^2$. The angle of incidence θ'_0 in the frame $\mathcal{R}'$ is related to θ by the relativistic aberration formulae:

$$\sin\theta'_0 = \frac{\sin\theta_0}{\gamma_0\left(1 - \beta_0\cos\theta_0\right)}, \quad \cos\theta'_0 = \frac{\cos\theta_0 - \beta_0}{1 - \beta_0\cos\theta_0}. \tag{8.2}$$

For very high Lorentz factors $\gamma_0 \gg 1$, this shows that $|\sin\theta'_0| \ll 1$ and $\cos\theta'_0 \simeq -1$, so that $\theta'_0 \simeq \pi$. Even if the photon field is close to isotropic, in the rest frame of sufficiently energetic electrons, the emission is strongly beamed in the electronic direction of motion. This is another occurrence of relativistic beaming. In the following, we will always assume

$\gamma_0 \gg 1$. Photons arriving with $\theta_0 \simeq 0$ have small energies in $\mathcal{R}'$ The largest photon energies are produced by head-on collisions $\theta_0 \simeq \pi$.

Inverse Compton scattering can actually be interpreted as direct Compton scattering in the electron rest frame. We can use the standard formula of direct Compton scattering (which can be proved using the conservation of 4-momentum), giving the energy of the scattered electron as a function of ϵ_0' and the angle between the scattered electron momentum and scattered photon momentum Θ':

$$\epsilon_1' = \frac{\epsilon_0'}{1 + \dfrac{\epsilon_0'}{m_e c^2}\left(1 - \cos\Theta'\right)}. \tag{8.3}$$

Since we assume $\gamma_0 \gg 1$, relativistic beaming implies that $\Theta' \simeq \pi - \theta_1'$ and $\cos\Theta' = -\cos\theta_1'$ (see Fig. 8.1). Now it is possible to express the energy of the scattered photon by transforming back into the observer frame, with $\beta_0 \simeq 1$:

$$\epsilon_1 = \gamma_0^2 \epsilon_0 \frac{\left(1 - \beta_0 \cos\theta_0\right)\left(1 + \beta_0 \cos\theta_1'\right)}{1 + \dfrac{\epsilon_0'}{m_e c^2}\left(1 + \cos\theta_1'\right)} \tag{8.4}$$

From these equations, two limiting cases can be defined.

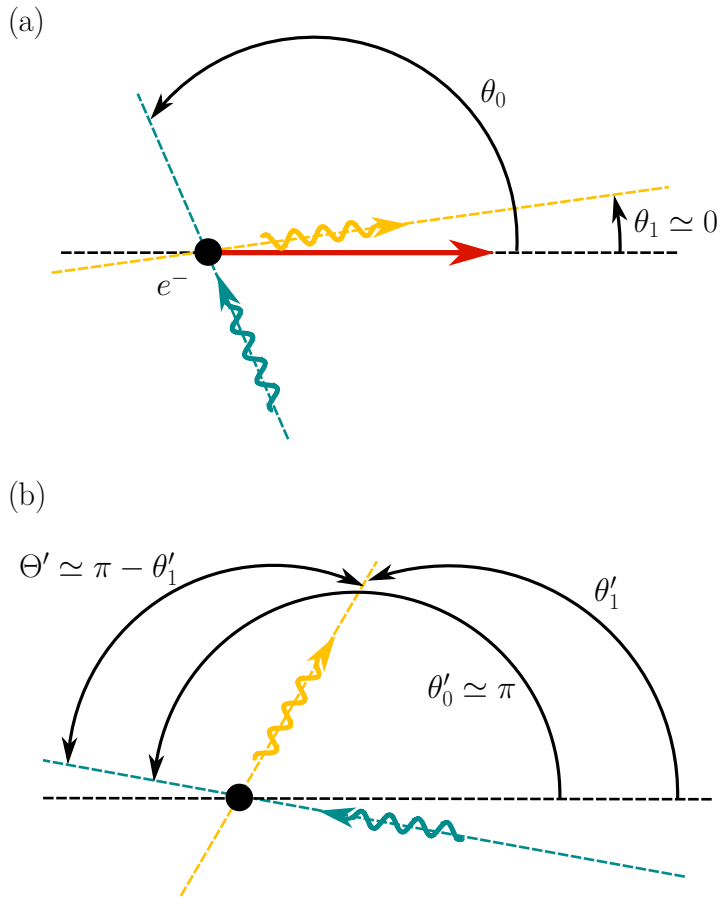


FIGURE 8.1: Scattering of a photon by an electron in the laboratory frame (a) and in the electron rest frame (b). The incident photon is pictured in blue, the scattered one in yellow. The red arrow shows the initial electron momentum. Angles are measured with respect to the initial direction of the electron.

- If the energy of the photon in the electron rest frame is much smaller than $m_e c^2$, then we have $\epsilon_1' \simeq \epsilon_0'$ from Eq. (8.3). The recoil of the electron is negligible, and the photon is scattered with practically the same energy (the collision is elastic). This is known as the Thomson regime.
- If the energy of the photon ϵ_0' is much larger than $m_e c^2$, the recoil of the electron cannot be ignored, and a consequent fraction of the electron energy is passed on to the photon (the collision is deeply inelastic). This is known as the Klein-Nishina regime.

In the Klein-Nishina regime $\gamma_0 \epsilon_0 \gg m_e c^2$, the maximal energy of the scattered photon is $\epsilon_{1,\max} \lesssim \gamma_0 m_e c^2$. In the Thomson regime $\gamma_0 \epsilon_0 \ll m_e c^2$, we know from Eq. (8.4) that the maximal energy is

$$\epsilon_{1,\max} \simeq 2\gamma_0 \epsilon_{1,\max}' \simeq 4\gamma_0^2 \epsilon_0. \tag{8.5}$$

This corresponds to a head-on collision, with $\theta_0 \simeq \pi$ and $\theta_1' \simeq 0$. In the Thomson regime, the characteristic energy of the scattered photon, although much larger than ϵ_0, is still small with respect to $\gamma_0 m_e c^2$.

Order of magnitude

The rest mass of the electron is $m_e c^2 = 511\,\mathrm{keV}$. This corresponds to a photon at the transition between hard X-rays and soft gamma rays, with frequency $\nu \simeq 10^{20}\,\mathrm{Hz}$ and a wavelength $\lambda \simeq 10^{-10}\,\mathrm{cm}$. For example, a photon in the near infra-red with energy $\epsilon_0 = 1\,\mathrm{eV}$, upscattered by an electron with Lorentz factor 10^4, would reach the gamma-ray band with a typical energy $\epsilon_1 \simeq 100\,\mathrm{MeV}$. This scattering would operate in the Thomson regime, since $\gamma_0 \epsilon_0 / m_e c^2 \simeq 0.02 \ll 1$.

8.1.2 Cross-section

The Thomson regime can be related to the classical scattering of an electromagnetic wave by an electron. The total cross-section of the Thomson scattering is constant, and given by

$$\sigma_\mathrm{T} = \frac{8\pi}{3} r_e^2 = 6.652 \times 10^{-25}\,\mathrm{cm}^2, \tag{8.6}$$

with $r_e = e^2/m_e c^2$ the classical radius of the electron. This does not depend on the polarization state of the incoming radiation. In this low-energy regime, one can afford to neglect the corrections from quantum electrodynamics (QED) that occur as $x = \epsilon_0'/m_e c^2$ approaches 1. The total cross-section of inverse Compton scattering in any regime is (Rybicki & Lightman, 1986)

$$\sigma_\mathrm{IC}(x) = \frac{3}{4}\sigma_\mathrm{T}\left(\frac{1+x}{x^3}\left(\frac{2x(1+x)}{1+2x} - \ln(1+2x)\right) + \frac{1}{2x}\ln(1+2x) - \frac{1+3x}{(1+2x)^2}\right). \tag{8.7}$$

It reduces to $\sigma_\mathrm{IC}(x) \simeq \sigma_\mathrm{T}$ for $x \ll 1$. In the Klein-Nishina regime, the cross section drops: interactions cost more energy to the electron but they occur less frequently. For $x \gg 1$, the cross section can be approximated by

$$\sigma_\mathrm{IC}(x) \simeq \frac{3}{8x}\sigma_\mathrm{T}\left(\ln(2x) + \frac{1}{2}\right). \tag{8.8}$$

The total cross section is sketched in Fig. 8.2.

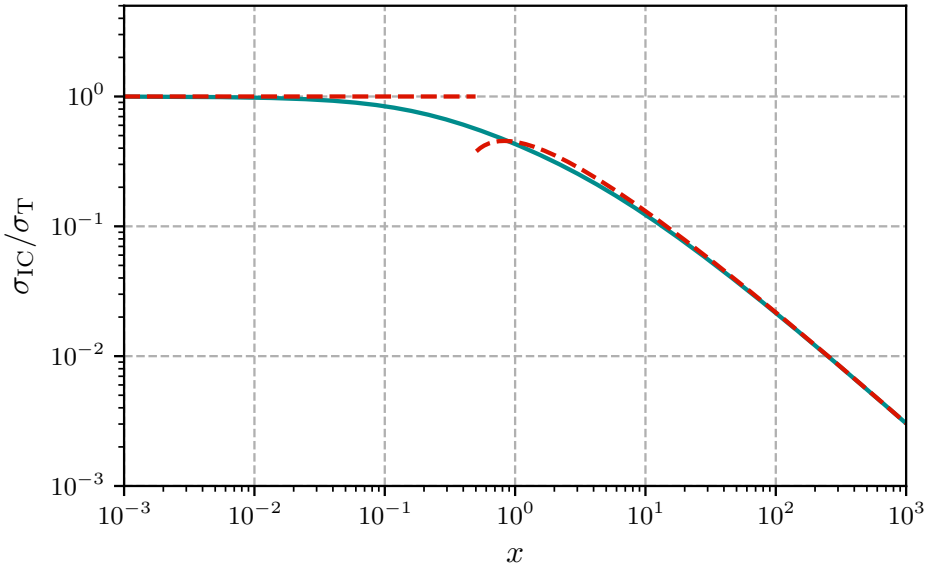


FIGURE 8.2: Cross section of the inverse Compton scattering process as a function of $x = \epsilon_0'/m_{\rm e}c^2$. At low x, it can be approximated with $\sigma_{\rm T}$, whereas Eq. (8.8) is a decent approximation at large x (dashed lines).

In the rest frame of the electron, the exact differential cross section is given by the Klein-Nishina formula:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_1'\,\mathrm{d}\epsilon_1'} = \frac{3}{16\pi}\sigma_{\rm T}\left(\frac{\epsilon_1'}{\epsilon_0'}\right)^2\left(\frac{\epsilon_1'}{\epsilon_0'} + \frac{\epsilon_0'}{\epsilon_1'} - \sin^2\Theta'\right), \tag{8.9}$$

where ϵ_0' is implicitly prescribed by Eq. (8.3) (this expression has the benefit of compactness). It is a measure of the probability that a photon should get scattered within the solid angle $\mathrm{d}\Omega_1'$, with energy in $[\epsilon_1', \epsilon_1' + \mathrm{d}\epsilon_1']$. In the Thomson regime, we have $\epsilon_0' \simeq \epsilon_1'$, so that the differential cross section reduces to

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_1'\,\mathrm{d}\epsilon_1'} \simeq \frac{3}{16\pi}\sigma_{\rm T}\left(1 + \cos^2\Theta'\right), \tag{8.10}$$

which is exactly the expression of the Thomson cross section of *unpolarized* radiation. There again, Klein-Nishina effects act mainly to reduce the probability of interaction. The radiation pattern of the differential cross section is represented in Fig. 8.3. At low energies, in the Thomson regime $\epsilon_0' \ll m_{\rm e}c^2$, the angular dependence is dipolar and matches that of the Thomson differential cross section (Eq. (8.10)). The scattering is symmetry with respect to Θ': forward and backward scatterings are equally likely. This symmetry is broken at higher energies, and forward scatterings ($\Theta' \simeq 0$) are favored.

8.1.3 Electron interacting with a photon field

Let us consider the interaction between a single electron and a photon field characterized by the photon number density $n(\epsilon)$: $n(\epsilon)\,\mathrm{d}\epsilon$ corresponds to the number of photons with energy in $[\epsilon, \epsilon + \mathrm{d}\epsilon]$. We assume that the photon field is isotropic in the laboratory frame. The cooling rate of the electron is $\mathrm{d}E/\,\mathrm{d}t$, with $E = \gamma_0 m_{\rm e}c^2$. Because time and energies transform the same way under a Lorentz boost, the cooling rate is a Lorentz invariant and we have $\mathrm{d}E/\,\mathrm{d}t = \mathrm{d}E'/\,\mathrm{d}t'$.

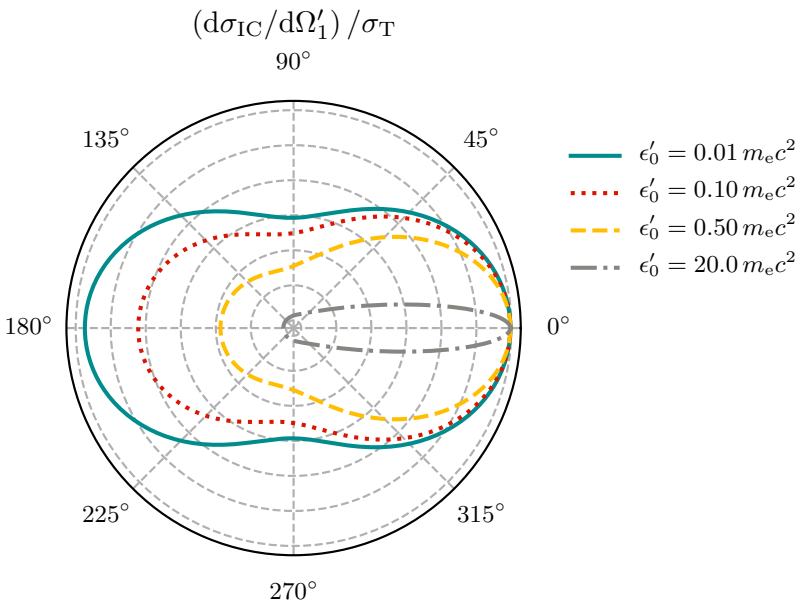


FIGURE 8.3: Angular dependence of the differential Klein-Nishina cross section, for different incident photon energies ϵ'_0. The polar angle is Θ'.

Thomson limit

In the Thomson limit, we know the expression of the radiated power measured in the rest frame of the electron, so that the cooling rate is

$$\frac{\mathrm{d}E'}{\mathrm{d}t'} = -\sigma_{\mathrm{T}} c \mathcal{E}', \tag{8.11}$$

where $\mathcal{E}' = \int \epsilon' \, \mathrm{d}n'$ is the total photon energy density in $\mathcal{R}'$. In the relativistic limit $\gamma_0 \gg 1$, if the photon field is isotropic, one can show that $\mathcal{E}'$ is related to $\mathcal{E}$ in the observer frame by $\mathcal{E}' = (4/3)\gamma_0^2 \mathcal{E}$ (Blumenthal & Gould, 1970). This yields the cooling rate of the electron

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{4}{3}\sigma_{\mathrm{T}} c \gamma_0^2 \mathcal{E}. \tag{8.12}$$

In the Thomson regime, the cooling rate is independent of the photon spectrum. The associated characteristic cooling time in the Thomson regime is

$$t_{\mathrm{IC}} = -\frac{E}{\mathrm{d}E/\,\mathrm{d}t} = -\frac{1}{\gamma_0}\frac{3 m_e c}{4\sigma_{\mathrm{T}}\mathcal{E}}. \tag{8.13}$$

From now on, let us assume for simplicity that the photon field is monoenergetic at energy ϵ_0, and has constant, uniform density n_0. In that casen $\mathcal{E}$ is simply $\mathcal{E} = n_0 \epsilon_0$. The subsequent results are easily generalized by replacing n_0 by $n(\epsilon_0)\,\mathrm{d}\epsilon_0$ and integrating on ϵ_0. The spectrum of photons that it produces is given by the following expression (Jones, 1968):

$$\frac{\mathrm{d}N_\gamma}{\mathrm{d}t\,\mathrm{d}\epsilon_1} = \frac{3}{4} c \sigma_{\mathrm{T}} \frac{n_0}{\gamma_0^2 \epsilon_0} \left(2x \ln(x) + x + 1 - 2x^2\right), \quad x = \frac{\epsilon_1}{4\gamma_0^2 \epsilon_0}. \tag{8.14}$$

This function is represented in Fig. 8.4 (the blue solid line for $\Gamma_\epsilon \ll 1$). In the Thomson limit, the maximal energy of a scattered photon is $4\gamma_0^2\epsilon_0$, so that x is contained in $[0, 1]$. The spetrum is peaked at the low-energy end.

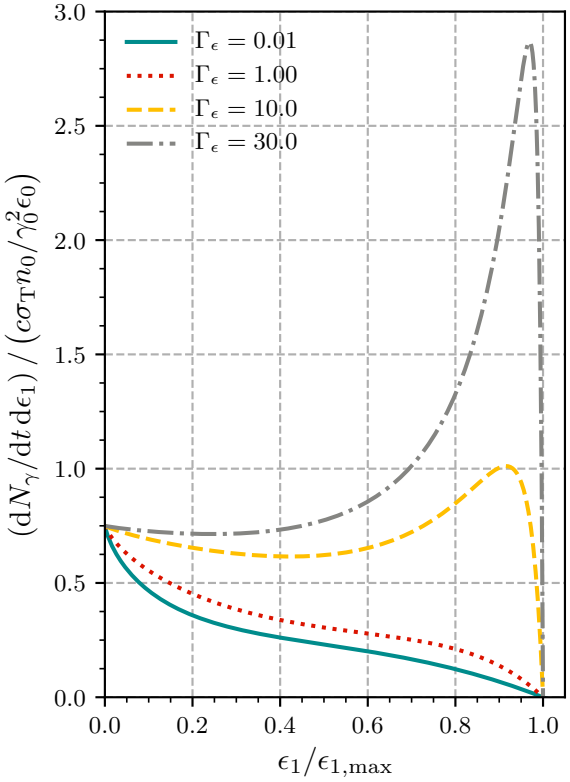


FIGURE 8.4: Spectrum of photons scattered by a single electron with Lorentz factor γ_0 traveling in a monoenergetic (at energy ϵ_0, isotropic and uniform bath. The spectrum, generally given by Eq. (8.17), depends on the parameter $\Gamma_\epsilon = 4\epsilon_0\gamma_0/m_\mathrm{e}c^2$. At low Γ_ϵ, the spectrum matches the prediction from the Thomson regime (Eq. (8.14)). Note that the energies on the x axis are normalized by the maximum energy $\epsilon_{1,\mathrm{max}}$, which increases as Γ_ϵ increases, allowing to show the different regimes on the same axis.

Klein-Nishina limit

Let us define the parameter $\Gamma_\epsilon = 4\epsilon_0\gamma_0/m_\mathrm{e}c^2$. The Thomson regime is characterized by $\Gamma_\epsilon \ll 1$, the Klein-Nishina regime by $\Gamma_\epsilon \gtrsim 1$. Let us rewrite Eq. (8.4) into the following form:

$$\epsilon_1 = \gamma_0 m_\mathrm{e}c^2 \frac{\Gamma_\epsilon\,(1-\beta_0\cos\theta_0)\,(1+\beta_0\cos\theta_1')\,/4}{1+\Gamma_\epsilon\,(1-\beta_0\cos\theta_0)\,(1+\beta_0\cos\theta_1')\,/4}. \tag{8.15}$$

Since $(1-\beta_0\cos\theta_0)\,(1+\beta_0\cos\theta_1')\,/4$ lies between 0 and 1, this equation implies that the maximal energy of a scattered photon is exactly

$$\epsilon_{1,\mathrm{max}} = \gamma_0 m_\mathrm{e}c^2 \frac{\Gamma_\epsilon}{1+\Gamma_\epsilon}. \tag{8.16}$$

As expected, this reduces to $\epsilon_{1,\mathrm{max}} \simeq 4\epsilon_0\gamma_0^2$ for $\Gamma_\epsilon \ll 1$ and $\epsilon_{1,\mathrm{max}} \simeq \gamma_0 m_\mathrm{e}c^2$ for $\Gamma_\epsilon \gg 1$. The spectrum of scattered photons, in the general case, was also computed by Jones (1968):

$$\frac{\mathrm{d}N_\gamma}{\mathrm{d}t\,\mathrm{d}\epsilon_1} = \frac{3}{4}c\sigma_\mathrm{T}\frac{n_0}{\gamma_0^2\epsilon_0}\left(2q\ln(q) + (1+2q)\,(1-q) + \frac{1}{2}\frac{(\Gamma_\epsilon)^2}{1+\Gamma_\epsilon}\,(1-q)\right), \tag{8.17}$$

where $q = \epsilon_1/\left(\Gamma_\epsilon\left(\gamma_0 m_\mathrm{e}c^2 - \epsilon_1\right)\right)$. For $\Gamma_\epsilon \ll 1$, this function approaches Eq. (8.14). For $\Gamma_\epsilon \gg 1$, the spectrum is strongly peaked near the maximal photon energy (see Fig. 8.4). This means that large energy losses are predominant.

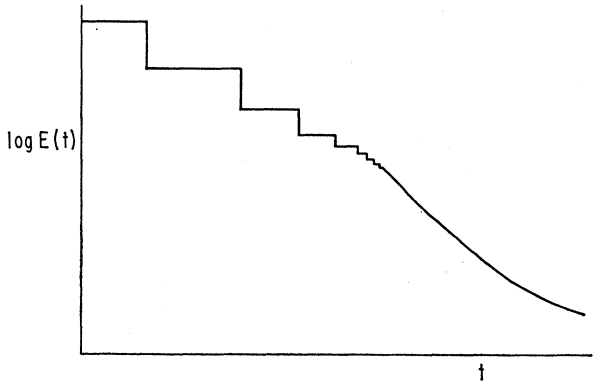


FIGURE 8.5: Typical evolution of the Lorentz factor of an energetic charged particle, first cooling in the Klein-Nishina regime, in logarithmic scale. Figure taken from Blumenthal & Gould (1970).

From Eq. (8.17), it is possible to derive a characteristic cooling rate, just like in the Thomson regime. However, since the scattering produces very high-energy photons in the Klein-Nishina regime, in general only a few encounters between the electron and surrounding photons are needed to cool the electron down. Its energy goes down by discrete steps, so a cooling rate is not particularly relevant. By contrast, because the cross section drops by a factor $\simeq \gamma_0 \epsilon_0 / m_\mathrm{e} c^2$, the mean free path of an ultra-relativistic electron increases by the same amount. The typical evolution of the energy of an electron which cools down in the Klein-Nishina regime, is shown in Fig. 8.5. At first, energy losses are discrete, occuring upon each scattering. Then the electron cools in the Thomson regime, where energy losses are more frequent and occur continuously.

Optical depth

By definition of the cross section, as the electron propagates along a distance $\mathrm{d}l$, the number of scattered photons is related to the scattering rate r by $\mathrm{d}N = r\,\mathrm{d}t$. Since the electron moves at a speed close to c, we have $\mathrm{d}N = r\,\mathrm{d}l/c$. Let us define τ the optical depth of the medium as the rate of absorption of the electron per unit of length l, that is,

$$\frac{\mathrm{d}\tau}{\mathrm{d}l} = \frac{r}{c}. \tag{8.18}$$

The optical depth can be physically interpreted as follows: the probability that the electron interacts with the photon field over a length $\mathrm{d}l$ is $\mathrm{d}\tau$, so that the probability that it does not interact is $1 - \mathrm{d}\tau$. Hence, over a macroscopic length L, the probability that the electron travels a distance L without being scattered is $\exp(-\tau)$, with $\tau = \int \mathrm{d}\tau = \int_0^L (\mathrm{d}\tau/\mathrm{d}l)\,\mathrm{d}l$.

Let us assume that the photon field is monoenergetic, but anisotropic in general. In order to compute the optical depth, we need to determine the scattering rate r. To do so, we move to the rest frame of the electron. By definition of the cross section, the scattering rate is

$$r' = \frac{\mathrm{d}N'}{\mathrm{d}t'} = c \int \sigma\,\mathrm{d}n', \tag{8.19}$$

where we integrate over all possible directions of incoming photons. To switch back to the observer frame, we use the result, proved by Blumenthal & Gould (1970), that $\mathrm{d}n/\epsilon$ is a Lorentz invariant, so that $\mathrm{d}n' = (\epsilon'/\epsilon)\,\mathrm{d}n$. Using the Lorentz boost of the energy $\epsilon_0' = \epsilon_0 \gamma_0 (1 - \beta_0 \cos\theta_0)$, and the time dilation $\mathrm{d}t = \gamma_0\,\mathrm{d}t'$, we finally obtain

$$r = \frac{1}{\gamma_0} r' = c \int (1 - \beta_0 \cos\theta_0)\, \sigma\left(\epsilon_0'\right) \frac{\mathrm{d}n}{\mathrm{d}\Omega_0}\, \mathrm{d}\Omega_0, \tag{8.20}$$

with $\Omega_0 = \sin\theta_0\, \mathrm{d}\theta_0\, \mathrm{d}\varphi_0$ the infinitesimal solid angle around a photon direction in the observer frame. Note that even if σ is the full cross section, it depends on θ_0 via the boosted energy energy ϵ_0'. Now, in the Thomson limit, σ is independent of ϵ_0' and θ_0. If n is isotropic ($\mathrm{d}n/\,\mathrm{d}\Omega_0 = n_0/\,(4\pi)$), then the $\cos\theta_0$ term integrates to zero, and we find the usual relation $r = c\sigma_\mathrm{T} n_0$.

> The difference with Eq. (8.20) comes essentially from the $1-\beta_0\cos\theta_0$ factor. Its physical meaning is that for relativistic electron speeds, this factor incorporates the fact that the relative velocity of the electron and the photon along the direction of the electron motion different than c.

From this, we can compute the optical depth of the electron traversing a distance L in the photon field:

$$\tau = \frac{n_0 \sigma_\mathrm{T}}{2} \int_{l=0}^{L} \int_{\theta_0=0}^{\pi} \frac{\sigma_\mathrm{IC}\left(\epsilon_0'\right)}{\sigma_\mathrm{T}} (1 - \beta_0 \cos\theta_0) \sin\theta_0\, \mathrm{d}\theta_0\, \mathrm{d}l. \tag{8.21}$$

In the context of black-hole magnetospheres, the characteristic optical depth of the medium over a characteristic length scale r_g is $\tau_0 = n_0 \sigma_\mathrm{T} r_\mathrm{g}$. It is customary to introduce the opacity $\kappa_\mathrm{IC}\left(\gamma_0\right)$, which depends on the Lorentz factor of the incoming electron, as

$$\tau = \int_0^L \kappa_\mathrm{IC}\, \mathrm{d}l, \quad \kappa_\mathrm{IC}\left(\gamma_0\right) = \frac{\tau_0}{2 r_\mathrm{g}} \int_0^\pi \frac{\sigma_\mathrm{IC}\left(\epsilon_0'\right)}{\sigma_\mathrm{T}} (1 - \beta_0 \cos\theta_0) \sin\theta_0\, \mathrm{d}\theta_0. \tag{8.22}$$

The optical depth, and the probability of interaction of the electron, depend on this Lorentz factor with the total cross section. The electron performs a random walk in the photon field, associated with a mean free path $\lambda \simeq 1/\left(\sigma_\mathrm{T} n_0\right)$ in the Thomson regime.

8.2 Photon-photon annihilation

8.2.1 Kinematics

Photon-photon annihilation describes the following interaction between two photons:

$$\gamma + \gamma \to e^+ + e^-. \tag{8.23}$$

Unlike Compton scattering, it cannot be described classically, even in a low-energy regime: it is entirely a QED process. For this process to happen, the energy of the photons must be of the order of the rest mass energy of an $e^\pm$ pair[1]. We can give a more precise sense to this statement.

Let us assume that the two photons, colliding with an angle θ_0, actually have enough energy to produce a pair, which has 4-momenta p_0, p_1 in the observer frame. The situation is pictured in Fig. 8.6. In this case, it is possible to define a center-of-mass (CM) frame, in which the produced pair has zero 3-momentum. Let us also denote k_0 and k_1 the 4-momenta of the two incident photons. The primed quantities will denote quantities measured in the CM frame. By conservation of the 4-momentum, we have $k_0^\mu + k_1^\mu = p_0^\mu + p_1^\mu$, so that $\left(k_0 + k_1\right)^2 = \left(p_0 + p_1\right)^2 = \left(p_0' + p_1'\right)^2$. Denoting ϵ_0 and ϵ_1 the energies of the photons in the observer frame, $E_-' = E_+' = E'$ the energies of the electron and positron in the CM frame, and $\boldsymbol{p}_-' = -\boldsymbol{p}_+'$

[1] This process can be seen as an extreme illustration of the mass-energy equivalence, describing how light alone can produce matter.

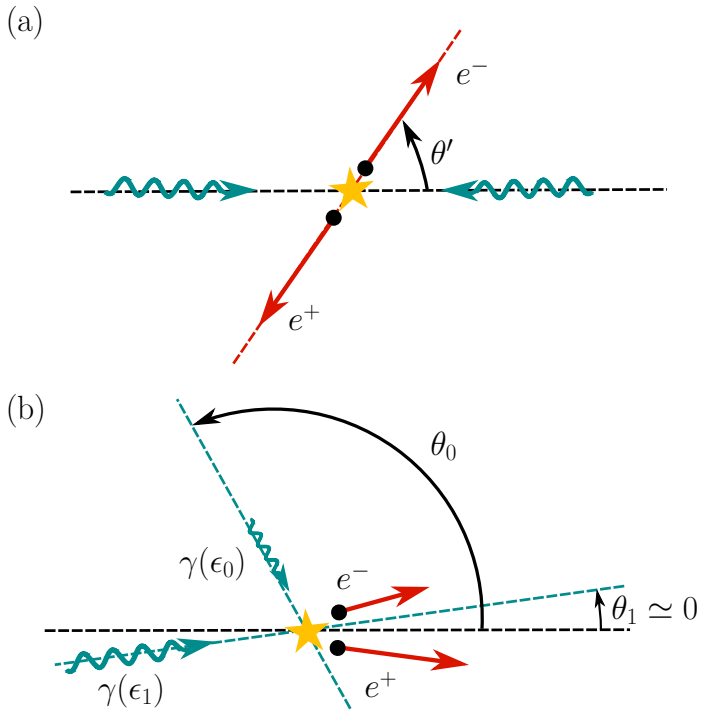


FIGURE 8.6: Production of a $e^\pm$ pair as seen from the center of mass frame (a) and the observer frame (b), in the case where the photon 1 is much more energetic than the other one ($\epsilon_1 \gg \epsilon_0$). The dashed black line marks the direction of the momentum of the center of mass, which is close to the direction of the momentum of photon 1.

the 3-momenta of the electron and positron in the CM frame, we obtain by developing the square:

$$2\epsilon_0\epsilon_1 \left(1 - \cos\theta_0\right) = 2m_\mathrm{e}^2c^4 + 2\left(E^2 + {\boldsymbol{p}'_-}^2\right). \tag{8.24}$$

The threshold for pair creation is reached when the produced leptons have zero momentum in the CM frame. This yields the following necessary condition for the annihilation of the two photons into a $e^\pm$ pair:

$$s = \frac{1}{2}\epsilon_0\epsilon_1 \left(1 - \cos\theta_0\right) \geq \left(m_\mathrm{e}c^2\right)^2. \tag{8.25}$$

This condition is all the easier to satisfy as θ_0 approaches π: a head-on collision is more likely to produce a pair, and the threshold condition reduces to $\epsilon_0\epsilon_1 \geq \left(m_\mathrm{e}c^2\right)^2$. The quantity s is actually the Lorentz invariant $s = \left(k_0 + k_1\right)^2/4$. We notice that $\left(k_0 + k_1\right)^2 = 4E'^2$, so that the Lorentz factor of the pair in the CM frame is

$$\gamma' = \frac{E'}{m_\mathrm{e}c^2} = \frac{\sqrt{s}}{m_\mathrm{e}c^2}. \tag{8.26}$$

Exactly at the pair creation threshold, we have $s = \left(m_\mathrm{e}c^2\right)^2$, so $\beta' = 0$ and $\gamma' = 1$.

> Even if a single photon in free space has enough energy to produce a pair, such event is unphysical because it would violate the conservation of momentum. However, it might occur in ultra-strong magnetic fields, which is a plausible pair plasma source in pulsar magnetospheres.

Eventually, we aim to reconstruct quantities in the observer frame. Hence, it is important to be able to transform from the observer frame to the CM frame, and vice versa. Let us determine the velocity $\beta_{\rm CM}$ (and the associated Lorentz factor $\gamma_{\rm CM}$) of the CM frame with respect to the observer frame (Cerutti, 2010). This is easier if we make the additional assumption that $\epsilon_1 \gg \epsilon_0$ (the second case in Sec. 8.2.2). In that case, the momentum of the photons is primarily carried by the photon 1, so that the momentum of the center of mass of the produced pair will be practically aligned with the momentum of photon 1 $\boldsymbol{p}_1$ (see Fig. 8.6). Consequently, the angle θ_1 between $\boldsymbol{p}_1$ and the momentum of the pair is negligible, and we can apply the following Lorentz boost to the energies of the incoming photons: $\epsilon_1' \simeq \gamma_{\rm CM}\,(1-\beta_{\rm CM})\,\epsilon_1$ and $\epsilon_0' \simeq \gamma_{\rm CM}\,(1-\beta_{\rm CM}\cos\theta_0)\,\epsilon_0$. At the same time, since the 3-momentum of the two photons vanishes in the CM frame, we know that $\epsilon_0' = \epsilon_1'$. Using the definition of s (Eq. (8.25)), this yields

$$\beta_{\rm CM} \simeq \frac{\epsilon_1 - \epsilon_0}{\epsilon_1 - \epsilon_0\cos\theta_0} \simeq 1 - \frac{2s}{{\epsilon_1}^2}; \quad \gamma_{\rm CM} = \frac{\epsilon_1}{2\sqrt{s}}. \tag{8.27}$$

8.2.2 Cross sections

The total cross section for the photon-photon annihilation, expressed as a function of the velocity of the pair in the CM frame $\beta' = \sqrt{1-1/\gamma'^2} = \sqrt{1-\left(m_{\rm e}c^2\right)^2/s}$, is given by Gould & Schréder (1967):

$$\sigma\left(\beta'\right) = \frac{3}{16}\sigma_{\rm T}\left(1-\beta'^2\right)\left(\left(3-\beta'^4\right)\ln\left(\frac{1+\beta'}{1-\beta'}\right) - 2\beta'\left(2-\beta'^2\right)\right). \tag{8.28}$$

The cross section is represented as a function of $s/(m_{\rm e}c^2)^2 = 1/(1-\beta'^2)$ in Fig. 8.7. As expected, it vanishes below the pair creation threshold $s = 1$. The cross section peaks at $\sigma_{\gamma\gamma} \simeq 0.25\,\sigma_{\rm T}$, and the peak is reached very close to the pair creation threshold. Assuming an isotropic photon field and averaging over θ_0, this means that on average pair creation occurs mostly when

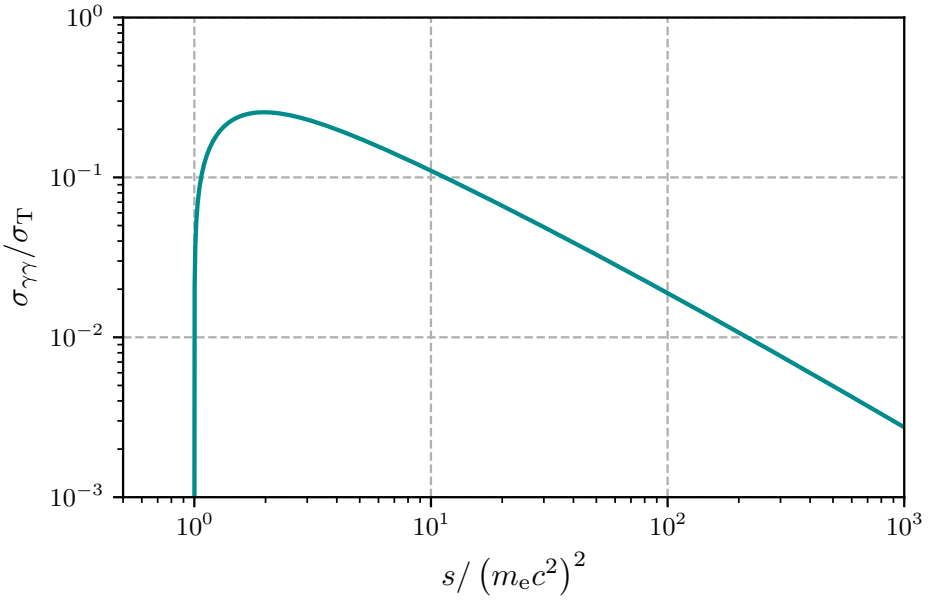


FIGURE 8.7: Cross section for pair creation by photon-photon annihilation as a function of $s/\left(m_{\rm e}c^2\right)^2$, where s is given by Eq. (8.25).

$$\epsilon_0 \epsilon_1 \simeq 4 m_\mathrm{e}^2 c^4. \tag{8.29}$$

Given Eq. (8.29), there are in general two situations in which photons can produce pairs. Either the two photons have similar energies $\epsilon_0 \simeq \epsilon_1 \simeq 2 m_\mathrm{e} c^2$, or one photon is much more energetic than the other ($\epsilon_1 \gg \epsilon_0$). In the following, we will make this approximation, which applies to our astrophysical context. This also ensures the validity of Eq. (8.27).

Order of magnitude

Given a photon with energy ϵ_0, the photons with which it is the most likely to interact have the energy

$$\epsilon_1 \simeq 1 \left(\frac{1\,\mathrm{TeV}}{\epsilon_0} \right) \mathrm{eV}. \tag{8.30}$$

In the first case, this means that there can be annihilation of two soft gamma ray photons, with energies of roughly $1\,\mathrm{MeV}$. In the second case, very high-energy photons with $\epsilon_1 \simeq 1\,\mathrm{TeV}$ can interact with soft near-infrared photons with $\epsilon_0 \simeq 1\,\mathrm{eV}$.

The differential cross section in the CM frame, as a function of β' and of the angle θ' between the produced pair and the incident photons in the CM frame (see Fig. 8.6), is given by Bonometto & Rees (1971)

$$\frac{\mathrm{d}\sigma_{\gamma\gamma}}{\mathrm{d}(\cos\theta')} = \frac{3}{16} \sigma_\mathrm{T} \beta' \left(1 - \beta'^2\right) \left(\frac{1 - (\beta' \cos\theta')^4 + 2\left(1 - \beta'^2\right) \beta'^2 \left(1 - \cos^2\theta'\right)}{\left(1 - (\beta' \cos\theta')^2\right)^2} \right). \tag{8.31}$$

Its radiation pattern is represented in Fig. 8.8. Close to the pair creation threshold ($\beta' \simeq 0$), the distribution is almost isotropic. On the other hand, as β' approaches 1, the differential cross sections peaks at $\theta' = 0$ and π. In that case, the pair is most likely emitted along the direction of the momentum of the center of mass, that is, the direction of the most energetic photon. The differential cross section is always symmetric with respect to $\cos\theta' = 0$.

Eq. (8.27) allows us to reconstruct the energies $E_\pm$ of the pair in the observer frame, using the backwards Lorentz transformation:

$$E_- = \gamma_\mathrm{CM} E' + \gamma_\mathrm{CM} \beta_\mathrm{CM} E' \beta' \cos\theta' = \gamma_\mathrm{CM} \left(\sqrt{s} + \beta_\mathrm{CM} \cos\theta' \sqrt{s - (m_\mathrm{e} c^2)^2} \right). \tag{8.32}$$

The energy of the positron is retrieved by the conservation of energy:

$$E_+ = \epsilon_1 + \epsilon_0 - E_- \simeq \epsilon_1 - E_-. \tag{8.33}$$

Note that although the electron and the positron have the same energies in the CM frame, they do not necessarily do in the observer frame because θ'_1 is a random variable. We chose to pick the energy of the electron first, but because the differential cross section is symmetric in $\cos\theta'$, there is no asymmetry between the two species. Indeed, for $\cos\theta' = 0$, we have $E_- = \gamma_\mathrm{CM} E' = \epsilon_1/2$, and $E_+ \simeq \epsilon_1 - E_- \simeq \epsilon_1/2 \simeq E_-$. The Lorentz factor of the CM frame is always much larger than 1 in the case $\epsilon_1 \gg \epsilon_0$. This implies that regardless of the angle θ', the electron and positron will be emitted in a direction very close to the momentum of the center of mass, because of relativistic beaming (this is pictured in Fig. 8.6).

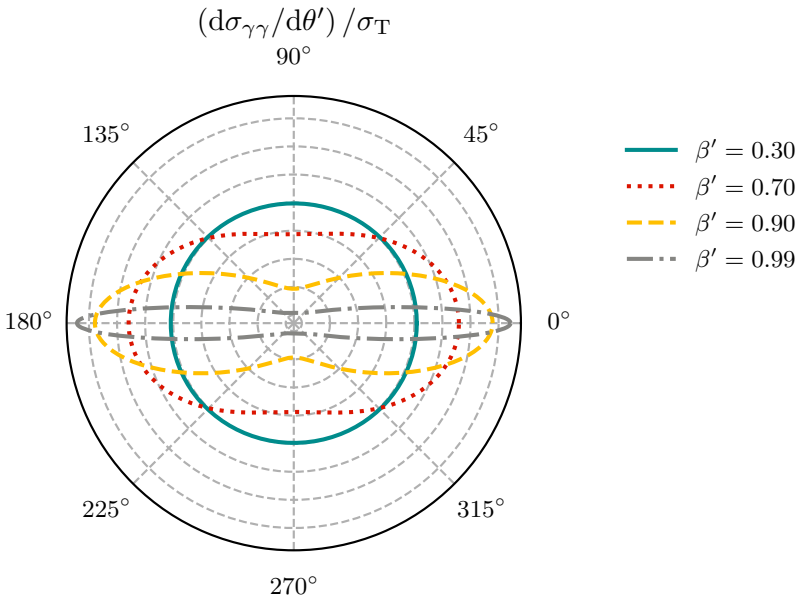


FIGURE 8.8: Angular dependence of the differential pair production cross section, for different velocities of the produced pair β' in the CM frame. The polar angle is the angle between the produced pair and the incoming photons in the CM frame θ'.

It seems that pair production has different properties depending on how far from the pair creation threshold the system is. To illustrate this, let us give the expression of the spectrum of electrons produced by an energetic photon with energy ϵ_1, propagating in an isotropic, monoenergetic and uniform photon field with energy ϵ_0 in the observer frame and densiy n_0, which reads (Aharonian et al., 1983; Zdziarski, 1988)

$$\frac{\mathrm{d}N_{e^-}}{\mathrm{d}t\,\mathrm{d}E_-} = \frac{3}{4}c\sigma_\mathrm{T}n_0\frac{1}{\zeta\epsilon_1}\left(\frac{1}{2}\left(\frac{E_-}{E_+}+\frac{E_+}{E_-}\right)-\left(2+\frac{1}{2}\left(\frac{E_-}{E_+}+\frac{E_+}{E_-}\right)\right)\frac{\chi}{\zeta}+2\left(\frac{\chi}{\zeta}\right)^2+2\frac{\chi}{\zeta}\ln\left(\frac{\zeta}{\chi}\right)\right), \tag{8.34}$$

where

$$\zeta = \frac{\epsilon_0\epsilon_1}{\left(m_\mathrm{e}c^2\right)^2}, \quad \chi = \frac{{\epsilon_1}^2}{4E_-E_+}, \quad E_+ = \epsilon_1 - E_-. \tag{8.35}$$

$E_+ = \epsilon_1 - E_-$ is the energy of the associated positron. This spectrum is represented in Fig. 8.9 for different values of the parameter ζ. As expected, the spectrum is invariant under the exchange between electron and positron energies $E_- \leftrightarrow E_+$. Close to the pair creation threshold ($\zeta \simeq 1$), the energy of the photon 1 is shared evenly between the electron and the positron ($E_- \simeq E_+ \simeq \epsilon_1/2$). As ζ increases, the energy of the photon tends to be mainly passed to one of the produced leptons ($E_+ \neq E_-$).

Finally, the optical depth for a high-energy photon propagating in this photon field can be computed in very much the same way as in the inverse Compton case:

$$\tau = \int_0^L \kappa_{\gamma\gamma}\,\mathrm{d}l, \quad \kappa_{\gamma\gamma}\left(\epsilon_1\right) \simeq \frac{\tau_0}{2r_\mathrm{g}}\int_0^\pi \frac{\sigma_{\gamma\gamma}\left(\beta'\right)}{\sigma_\mathrm{T}}\left(1-\cos\theta_0\right)\sin\theta_0\,\mathrm{d}\theta_0. \tag{8.36}$$

The integration is carried over all possible collision angles θ_0 between the soft photon 0 and the high-energy photon 1. We also assumed $\beta_\mathrm{CM} \simeq 1$ for simplicity, which is valid if $\epsilon_1 \gg \epsilon_0$.

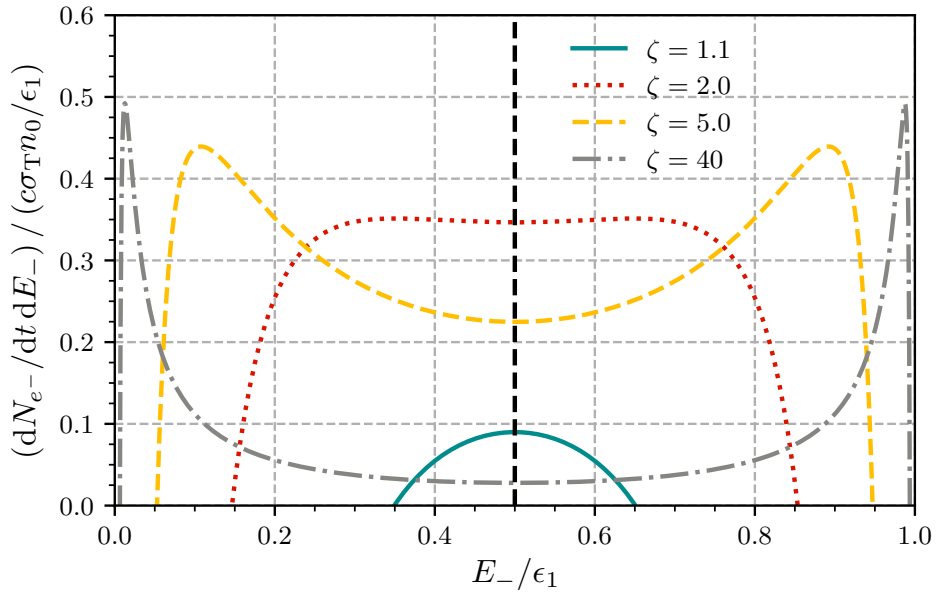


FIGURE 8.9: Energy spectrum of the electrons produced by a high-energy photon of energy ϵ_1 propagating in a uniform, isotropic and monoenergetic (with energy ϵ_0) bath of soft photons, in the limit $\epsilon_1 \gg \epsilon_0$. The spectrum depends on the parameter $\zeta = \epsilon_0\epsilon_1/(m_\mathrm{e}c^2)^2$. The electron energy is bounded by ϵ_1.

In the following of this thesis, we will not be concerned with the inverse process: the pair recombination $e^+ + e^- \to \gamma + \gamma$, despite the fact that this process is not limited by a threshold. This is because we will assume very energetic leptons with Lorentz factors much larger than 1. In this ultra-relativistic limit, the pair annihiliation cross section is very well approximated by (Dirac, 1930; Svensson, 1982)

$$\sigma_{e^+e^-} \simeq \frac{3}{8}\sigma_\mathrm{T}\frac{\ln(2\gamma_0) - 1}{\gamma_0}, \tag{8.37}$$

where γ_0 is the Lorentz factor of the pair in the CM frame. The cross section drops considerably for $\gamma_0 \gg 1$, and only the direct process is relevant in our context.

8.3 Synchrotron radiation

Synchrotron radiation is emitted by relativistic electrons or positrons spiralling around magnetic field lines. We will review some useful results, especially in the case where the emitting particles are also submitted to an electric field. Most of the material covered in this section can be found in much more detail in Ginzburg & Syrovatskii (1965), Blumenthal & Gould (1970) and Rybicki & Lightman (1986). We will simply use a classical treatment of the synchrotron radiation, without including QED effects.

8.3.1 Emitted power

In a uniform magnetic field $\boldsymbol{B} = B\boldsymbol{e}_z$, with zero electric field, an electron moves in a spiral path at the gyrofrequency

$$\nu_\mathrm{g} = \frac{eB}{2\pi c m_\mathrm{e}\gamma}, \tag{8.38}$$

with γ the Lorentz factor of the particle. Let us consider a trajectory with a pitch angle α between the velocity of the particle and the magnetic field. The associated Larmor radius

is $r_{\rm L} = v_\perp/(2\pi\nu_{\rm g}) = \gamma m_{\rm e}c^2\beta\sin\alpha/(eB)$. In the non-relativistic limit, the gyrofrequency reduces to the Larmor frequency $\nu_{\rm c} = eB/(2\pi c m_{\rm e})$. In this limit, the usual Larmor formula gives the radiation loss rate:

$$P_{\rm Larmor} = \frac{2e^2\boldsymbol{a}^2}{3c^3}, \tag{8.39}$$

where $\boldsymbol{a}$ is the 3-acceleration of the particle. A covariant generalization for a particle of 4-momentum p^μ is given by

$$P_{\rm rad} = \frac{2e^2}{3m_{\rm e}^2c^3}\frac{{\rm d}p_\mu}{{\rm d}\tau}\frac{{\rm d}p^\mu}{{\rm d}\tau}. \tag{8.40}$$

This is independent of the underlying acceleration mechanism. From Eq. (8.40), one can prove that a particle linearly accelerated will radiate less than a particle in uniform circular motion by a factor γ^2, for a given magnitude of applied force (Jackson, 1998). As a result, if the particle is energetic enough, it will be assumed to move instantaneously along a circular path. The covariant Larmor formula for such Larmor gyrations yields

$$P_{\rm rad} = \frac{2e^4}{3m_{\rm e}^2c^3}\beta^2\gamma^2B^2\sin^2\alpha = \frac{1}{4\pi}c\sigma_{\rm T}\beta^2\gamma^2B_\perp^2, \tag{8.41}$$

where we defined the perpendicular magnetic field $B_\perp = B\sin\alpha$. This formula is valid for a fixed pitch angle. For a large population of high-energy particles, the distribution of pitch angles is likely to be isotropic. These particle have a characteristic cooling time scale $t_{\rm sync}$, which is obtained by averaging Eq. (8.41) over the pitch angles θ. With $\langle\sin^2\theta\rangle = 2/3$, and introducing $\mathcal{U}_B = B^2/(8\pi)$ as the volume density of electromagnetic energy, we find

$$\langle P_{\rm rad}\rangle = \frac{4}{3}c\sigma_{\rm T}\mathcal{U}_B\beta^2\gamma^2 \Rightarrow t_{\rm sync} = \frac{\gamma m_{\rm e}c^2}{P_{\rm rad}} = \frac{1}{\gamma}\frac{3m_{\rm e}c}{4\sigma_{\rm T}\mathcal{U}_B}. \tag{8.42}$$

We notice that this expression is extremely similar to the cooling time by inverse Compton scattering, where we replaced the radiation energy density by the magnetic energy density[2].

> The dependence of the radiated power and the cooling time on the mass of the particle for both inverse Compton scattering and synchrotron radiation shows that electrons cool much more efficiently than protons.

8.3.2 Effective perpendicular magnetic field

If the particle propagates in both a magnetic and an electric field, Eq. (8.40) actually turns into

$$P_{\rm tot} = \frac{1}{4\pi}c\sigma_{\rm T}\beta^2\gamma^2\left((\boldsymbol{E}+\boldsymbol{\beta}\times\boldsymbol{B})^2 - (\boldsymbol{\beta}\cdot\boldsymbol{E})^2\right). \tag{8.43}$$

This equation reduces to Eq. (8.41) if $\boldsymbol{E} = \boldsymbol{0}$. If $\boldsymbol{B} = 0$, we find the power radiated by an electron lineraly accelerated by an electric field, which is indeed reduced by a factor γ^2. In most cases, the term $\boldsymbol{\beta}\times\boldsymbol{B}$ dominates. The similarity between Eq. (8.41) and Eq. (8.43) can be made more apparent by defining the field (Cerutti et al., 2016)

$$\tilde{\boldsymbol{B}}_\perp = \boldsymbol{E} + \boldsymbol{\beta}\times\boldsymbol{B} - (\boldsymbol{\beta}\cdot\boldsymbol{E})\,\boldsymbol{\beta}. \tag{8.44}$$

One can check that $\boldsymbol{\beta}\cdot\tilde{\boldsymbol{B}}_\perp = \boldsymbol{0}$ and $\tilde{\boldsymbol{B}}_\perp^2 = (\boldsymbol{E}+\boldsymbol{\beta}\times\boldsymbol{B})^2 - (\boldsymbol{\beta}\cdot\boldsymbol{E})^2$ in the relativistic limit $\boldsymbol{\beta}^2 \simeq 1$. This allows us to rewrite Eq. (8.43) as

[2]Synchrotron radiation can be viewed as inverse Compton scattering by the spiralling electron of virtual photons associated with the magnetic field (Blumenthal & Gould, 1970).

$$P_{\rm tot} = \frac{1}{4\pi} c\sigma_{\rm T}\beta^2\gamma^2\tilde{B}_\perp^2, \quad \tilde{B}_\perp = \sqrt{\tilde{\boldsymbol{B}}_\perp^2} \tag{8.45}$$

When an electric field is present, the perpendicular magnetic field in Eq. (8.41) must simply be replaced by $\tilde{B}_\perp$. This allows us to interpret $\tilde{\boldsymbol{B}}_\perp$ as the effective perpendicular magnetic field. Another illuminating way to understand this effective field is to look at the equations of motion of a particle with velocity $c\boldsymbol{\beta}$ (Kelner et al., 2015):

$$\frac{\mathrm{d}(\gamma\boldsymbol{\beta})}{\mathrm{d}t} = \frac{q}{m_{\rm e}c}\left(\boldsymbol{E} + \boldsymbol{\beta}\times\boldsymbol{B}\right) \tag{8.46}$$

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \frac{q}{m_{\rm e}c}\left(\boldsymbol{E}\cdot\boldsymbol{\beta}\right). \tag{8.47}$$

Expanding the left-hand side of Eq. (8.46) and inserting Eq. (8.47) yields

$$\frac{\mathrm{d}\boldsymbol{\beta}}{\mathrm{d}t} = \frac{q}{\gamma m_{\rm e}c}\left(\boldsymbol{E} + \boldsymbol{\beta}\times\boldsymbol{B} - (\boldsymbol{\beta}\cdot\boldsymbol{E})\,\boldsymbol{\beta}\right), \tag{8.48}$$

while in the absence of electric field, since the Lorentz factor of the particle remains constant, the equation of motion reduces to $\mathrm{d}\boldsymbol{\beta}/\,\mathrm{d}t = (\boldsymbol{\beta}\times\boldsymbol{B})/(\gamma m_{\rm e}c)$. This means that $\tilde{\boldsymbol{B}}_\perp$ is indeed the perpendicular magnetic field in the general case.

8.3.3 Spectrum emitted by a gyrating particle

As mentioned earlier, we focus on the case of a particle having its velocity and acceleration perpendicular. In the rest frame of the particle, the radiation pattern of the accelerated charge is dipolar: we have

$$\frac{\mathrm{d}P'_{\rm rad}}{\mathrm{d}\Omega'} = \frac{e^2\boldsymbol{a}'^2}{4\pi c^3}\sin^2\Theta', \tag{8.49}$$

where the primed quantities are measured in the rest frame of the particle and where Θ' is the angle between its acceleration and the direction of emission. On the other hand, the received radiated power is strongly beamed toward its direction of motion in the observer frame, due to relativistic effects. This is illustrated in Fig. 8.10. A major part of the energy is emitted in a cone with semi-aperture angle $1/\gamma \ll 1$. This has dramatic consequences concerning the spectrum of synchrotron radiation measured by an observer at rest in the observer frame.

Let us consider how the emitted radiation and its spectrum evolve as the Lorentz factor of the particle γ increases. At low energies, the received intensity varies sinusoidally, precisely because the radiation pattern of the particle is dipolar. The spectrum consists of a single line at the Larmor frequency $\nu_{\rm c} = qB/(2\pi m_{\rm e}c)$: this is often named cyclotron radiation. As γ increases, higher harmonics of the fundamental frequency $\nu_{\rm g} = \nu_{\rm c}/\gamma$ of the motion appear in the spectrum, because relativistic beaming distorts the observed angular distribution of the intensity from a $\sin^2\theta$ law. At relativistic velocities, the spectrum involves a large number of harmonics which contribute significantly to the emitted power. The spectral lines broaden due to Doppler effects, aberration effects, or because there is a spread in the energies of the electron population. The broadening being proportional to the emission frequency, except at frequencies close to the fundamental, the emission spectrum becomes continuous (see Fig. 8.11(a)). The envelope of all these harmonics is the spectrum. Details on cyclo-synchrotron spectrum can be found in Bekefi (1966). In the ultra-relativistic limit $\gamma \gg 1$, the enveloppe of the spectral power reads

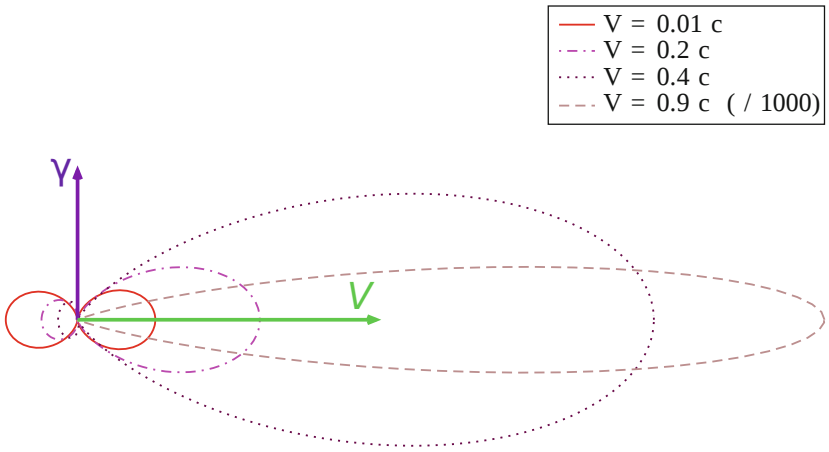


FIGURE 8.10: Radiation pattern of an accelerated charge with velocity $\boldsymbol{V}$ orthogonal to its acceleration. The diagram for $V = 0.9c$ was rescaled by a factor $1/1000$ to fit in the figure. The emission cone gets narrower as V approaches c. Figure taken from Gourgoulhon (2010).

(a) (b)

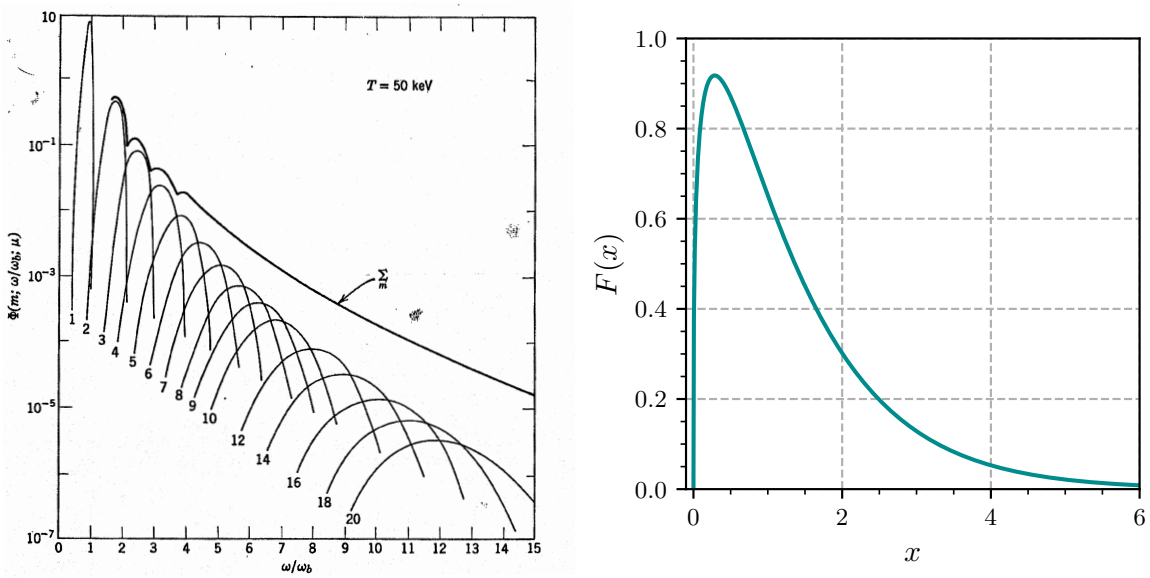


FIGURE 8.11: (a) Spectrum of the first twenty synchrotron harmonics and cumulative spectrum from mildly relativistic charged particles. Figure taken from Bekefi (1966). (b) Synchrotron kernel $F(x)$, given in Eq. (8.50).

$$\frac{\mathrm{d}P_{\rm rad}}{\mathrm{d}\nu} = \frac{\sqrt{3}e^3 B_\perp}{m_{\rm e}c^2} F\left(\frac{\nu}{\nu_{\rm s}}\right), \quad F(x) = x\int_x^{+\infty} K_{5/3}(y)\,\mathrm{d}y, \quad \nu_{\rm s} = \frac{3eB_\perp\gamma^2}{4\pi m_{\rm e}c}. \tag{8.50}$$

The graph of the function F is shown in Fig. 8.11(b). Importantly, the characteristic frequency of the spectrum is $\nu_{\rm s} \simeq \gamma^3\nu_{\rm g} \gg \nu_{\rm g}$. The distribution is peaked at $\nu \simeq 0.29\,\nu_{\rm s}$. Most of the power is emitted at $\nu \gg \nu_{\rm g}$, so that the spectrum can be considered smooth and continuous. In the case where an electric field is present, one simply needs to replace $B_\perp$ by $\tilde{B}_\perp$ in Eq. (8.50).

Order of magnitude

The numerical value of the characteristic synchrotron frequency is

$$\nu_\mathrm{s} = 4.2 \left(\frac{B_\perp}{1\,\mathrm{G}}\right) \gamma^2 \mathrm{MHz}. \tag{8.51}$$

For field strengths $B_\perp \simeq 10^2$ G, the synchrotron emission is in the radio-mm band for moderately relativistic electrons.

8.3.4 Radiation reaction force

Since the motion of charged particles perforce involves the emission of radiation, carrying energy and momentum, the radiation must have a feedback on the motion of the particle. This self-consistent problem is in general very difficult to solve: radiated electromagnetic fields brake the particle, which affects the radiation itself. The equation of motion of the particle is in general

$$\frac{\mathrm{d}u^\mu}{\mathrm{d}\tau} = \frac{e}{m_\mathrm{e}c} F^{\mu\nu} u_\nu + g^\mu, \tag{8.52}$$

where $F^{\mu\nu}$ is the Maxwell tensor, u^μ the 4-velocity of the particle and g^μ the sought reaction force.

In the rest frame of the particle, the reaction force must match the Abraham-Lorentz force $\boldsymbol{F} = \left(2q^2/3c\right) \ddot{\boldsymbol{\beta}}$ (Jackson, 1998; Landau & Lifshitz, 1980). This expression of the force leads to contradictions in the absence of external electromagnetic fields, as the free particle could increase its own energy. Actually, this expression for $\boldsymbol{F}$ is only valid in the limit where the radiation force is small with respect to the external Lorentz force. Denoting λ the wavelength of the fields and B their characteristic strength, this leads to the self-consistency conditions

$$B \ll B_\mathrm{c} = \frac{m_\mathrm{e}^2 c^4}{e^3}, \quad \lambda \gg \frac{e^2}{m_\mathrm{e}c^2} = r_\mathrm{e}. \tag{8.53}$$

In the rest frame of the particle, classical electrodynamics break down when the magnetic field strength exceeds the "critical" magnetic field $B_\mathrm{c} = 6.0 \times 10^{15}$ G, or when distances shorter than the classical radius of the electron are probed by the fields. When B approaches B_c, the Larmor radius $r_\mathrm{L} = m_\mathrm{e}c^2/eB$ of a particle approaches the classical radius r_e. Beyond this regime, classical electrodynamics are inconsistent.

> Actually, QED effects become important for magnetic field strengths above $B_\mathrm{QED} = m_\mathrm{e}^2 c^3/e\hbar = 4.4 \times 10^{13}$ G. When B approaches B_QED, the Larmor radius of the particle approaches the Compton wavelength $\lambda_\mathrm{c} = h/m_\mathrm{e}c$, and the typical energy of a synchrotron photon $h\nu_\mathrm{s} \simeq \gamma^2 h\nu_\mathrm{c}$ is similar to the particle energy $\gamma m_\mathrm{e}c^2$. In some way, since $B_\mathrm{QED} \ll B_\mathrm{c}$, quantum effects protect classical electrodynamics from its own inconsistencies. In the quantum regime, just like in the Klein-Nishina regime, synchrotron photons are emitted in discrete steps and carry a sizeable portion of the particle energy.

Landau & Lifshitz (1980) derived a covariant generalization of the previous results, valid in any frame, in the relativistic case. A general expression for the radiation force is:

$$g^\mu = \frac{2q^2}{3c^3} \frac{\mathrm{d}^2 u^\mu}{\mathrm{d}\tau^2} - \frac{P_\mathrm{tot}}{c^2} u^\mu, \tag{8.54}$$

which satisfies the requirements that $g^\mu u_\mu = 0$, that g^i reduces to F^i in the non-relativistic limit, and that it gives the correct radiated power. The associated 3-force is

$$\boldsymbol{g} = \frac{2q^3}{3m_\mathrm{e}c^3}\gamma\left(\left(\partial_t + c\boldsymbol{\beta}\cdot\boldsymbol{\nabla}\right)\boldsymbol{E} + \boldsymbol{\beta}\times\left(\partial_t + c\boldsymbol{\beta}\cdot\boldsymbol{\nabla}\right)\boldsymbol{B}\right) + \frac{\sigma_\mathrm{T}}{4\pi}\left(\left(\boldsymbol{E}+\boldsymbol{\beta}\times\boldsymbol{B}\right)\times\boldsymbol{B} + \left(\boldsymbol{\beta}\cdot\boldsymbol{E}\right)\boldsymbol{E}\right) - \frac{\sigma_\mathrm{T}}{4\pi}\gamma^2\left(\left(\boldsymbol{E}+\boldsymbol{\beta}\times\boldsymbol{B}\right)^2 - \left(\boldsymbol{\beta}\cdot\boldsymbol{E}\right)^2\right)\boldsymbol{\beta}. \tag{8.55}$$

In the following, we will be concerned with relativistic particles with $\gamma \gg 1$. In that case, the last term is the leading term (although the second term can be necessary to correctly include curvature radiation; see Cerutti et al. 2016). The radiation reaction force reduces to

$$\boldsymbol{g} \simeq -\frac{\sigma_\mathrm{T}}{4\pi}\gamma^2\left(\left(\boldsymbol{E}+\boldsymbol{\beta}\times\boldsymbol{B}\right)^2 - \left(\boldsymbol{\beta}\cdot\boldsymbol{E}\right)^2\right)\boldsymbol{\beta} = -\frac{P_\mathrm{tot}}{c}\boldsymbol{\beta}. \tag{8.56}$$

The reaction force indeed acts like a drag in the direction opposite to the particle velocity.

We have pointed out that this procedure is only applicable if the strength of the fields in the rest frame of the particle are small with respect to B_c. Let us denote B the typical strength of the fields in the observer frame: this means that the fields have the strength γB in the rest frame of the particle. Therefore the domain of validity of Eq. (8.56) is (Landau & Lifshitz, 1980; Blumenthal & Gould, 1970; Cerutti et al., 2012)

$$\gamma B \ll B_\mathrm{c}. \tag{8.57}$$

This is the condition under which the reaction force is negligible in the particle rest frame. Interestingly, even if this condition is satisfied, the magnitude of the 3-force can exceed the Lorentz force in the observer frame, because the ratio of the two is roughly $\gamma^2\left(B/B_\mathrm{c}\right)$.

Part III
Numerical techniques

"M. Fourier avait l'opinion que le but principal des mathématiques était l'utilité publique et l'explication des phénomènes naturels; mais un philosophe comme lui aurait dû savoir que le but unique de la science, c'est l'honneur de l'esprit humain, et que sous ce titre, une question de nombres vaut autant qu'une question du système du monde."

"Mr. Fourier had the opinion that that the main purpose of mathematics was public utility and explanation of natural phenomena; but a philosopher of such distinction should have known that the sole end of science is the honor of the human mind, and that under this title a question about numbers is worth as much as a question about the system of the world."

CARL GUSTAV JAKOB JACOBI, lettre à André-Marie Legendre, 2 juillet 1830

Chapter 9

Particle-in-cell simulations

We have exhibited in Part. II a large number of phenomena: general relativity, electrodynamics, plasma behaviour, radiative processes. It looks hopeless to treat this intricate and nonlinear problem other than numerically. In this section, I will describe a specific numerical technique designed for numerical simulations of plasmas: the particle-in-cell (hereafter, PIC) method. Although the basic laws governing the plasma are Newton's and Maxwell's equations, solving them for a macroscopic number of particles is extremely challenging. The principles of special relativistic PIC methods will be presented for simple geometries. Some excellent material on PIC codes and plasma simulations can be found in the textbooks by Birdsall & Langdon (1985) and Hockney & Eastwood (1988), to which the reader is referred for advanced aspects. The numerical aspects of general relativity are deferred to Chap. 10.

> There is a great variety of PIC algorithms, for almost as many simulations methods (implicit/explicit integration schemes, grid interpolation procedures,...). In this chapter, we only describe and justify the choices made in the `Zeltron` code; other choices make perfect sense in other contexts.

9.1 Principles

In the introduction, we detailed why traditional MHD methods are ill-advised in the context of this thesis. One of the prime reasons is that we aim to simulate a collisionless plasma. Such a plasma is most generally characterized by the distribution function $f_s(\boldsymbol{x}, \boldsymbol{p}, t)$ (for each species s), the evolution of which is ruled by Vlasov's equation:

$$\frac{\partial f_s}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} f_s + q_s(\boldsymbol{E} + \frac{\boldsymbol{v}}{c} \times \boldsymbol{B}) \cdot \boldsymbol{\nabla}_{\boldsymbol{p}} f_s = 0, \tag{9.1}$$

with $\boldsymbol{p} = \gamma m_s \boldsymbol{v}$ the momentum of the particle, q_s and m_s the charge and mass of the species s. These equations must be solved concurrently with Maxwell's equations. It is possible to solve directly the Vlasov equation, treating the distribution in phase space as a fluid and employing hydrodynamic solvers. The upside is the very low level of noise, which is convenient in order to capture low-amplitude effects, the linear growth of instabilities, or broad distribution functions. In addition, important inhomogeneities in density cause no issue. On the other hand, these solvers are very demanding computationally, because the Vlasov equation is 6-dimensional. See Palmroth et al. (2018) for a review, or Umeda et al. (2009) for an example of implementation.

Straightforward integration the Vlasov equation is often too challenging. The number of degrees of freedom can be reduced through the gyrokinetic approximation, which can be applied to highly magnetized plasmas, with slow dynamics compared to the cyclotron motion with small departure from the equilibrium (Howes et al., 2006). The idea is to average the Vlasov equation over the Larmor gyrations of the particles. Relaxing the constraint of resolving Larmor orbits is very beneficial in the case of extremely magnetized plasmas.

These codes have long been used in the context of fusion plasmas, but they have also been recently applied to high-energy astrophysical problems (Numata et al., 2010).

Another idea is to solve the Vlasov equation indirectly. In the particle approximation, the plasma is treated as a set of charged macroparticles. Integrating the discrete particle trajectories can be shown to be equivalent to solving the full kinetic collisionless Vlasov equation along characteristic curves (Sironi & Cerutti, 2017). In this approach, the distribution function for N macroparticles with positions $\boldsymbol{r}_k(t)$ and momenta $\boldsymbol{p}_k(t)$ is approximated by

$$f_s(\boldsymbol{r}, \boldsymbol{p}, t) = \sum_{k=1}^{N} w_k \delta(\boldsymbol{r} - \boldsymbol{r}_k(t)) \delta(\boldsymbol{p} - \boldsymbol{p}_k(t)). \tag{9.2}$$

In Eq. (9.2), the coefficient w_k is called the weight of the particle. This prefactor aims at compensating for the low number of numerical particles (at most 10^{12} in large simulations) with respect to the actual number of particles in the plasma (from 10^6 to 10^{10} particles per Debye sphere, Fitzpatrick, 2015). Using a large weight, numerical macroparticles represent a very large number of physical particles with the same charge-to-mass ratio, so they follow the same phase space trajectories. The main drawback of this approach comes with its discrete nature: the phase space is much harder to resolve, which leads to shot noise.

Even though the plasma is collisionless, macroparticles feel long-range binary interactions, and a brute force approach would require the computation of $N(N-1)/2 \propto N^2$ interactions, which would be too expensive. Instead, the electromagnetic fields are computed on a mesh, and then interpolated back to the particles. This is the heart of the PIC method (a "particle-mesh" method), which reduces the number of operations per time step to $\sim N \log N$ (Pritchett, 2003), and eliminates fluctuations on a scale smaller than the grid spacing. Not only is this convenient numerically, but this is also a defining feature of collisionless plasma physics: the dynamics of particles are driven by collective processes, operating through the electromagnetic fields, rather than by individual collisions which are not resolved by the PIC method. Incidentally, this also removes the matter of the singularities in Coulomb's law: plasma physics are only relevant at scales larger than the Debye length. In PIC simulations, resolving the microscopic kinetic scales is absolutely necessary.

> The scale of the downsampling needed to go from physical plasmas to numerical plasmas is so large that one may question the relevance of plasma simulations. Birdsall & Langdon (1985) and Hockney & Eastwood (1988) showed that for PIC algorithms, the plasma parameter Λ (the number of particles per Debye sphere) need actually be only of a few.

Most plasma simulations simulate an electron/ion plasma, which are characterized by a large mass ratio between the two species. Some hybrid models have been developed to retain some aspects of collisionless plasma physics, while gaining orders of magnitude in resolution or in the duration of the simulation by adding some fluid description. Most of the time, the physics at the electronic scale is discarded, whereas ions are still described kinetically. In the hybrid-PIC method for instance, electrons are considered massless, so they constitute a perfectly conducting fluid (Winske et al., 2003). An MHD-PIC approach has also been developed to treat kinetically one species (e.g., cosmic rays), and study the interaction between this collisionless plasma and a thermal plasma (Bai et al., 2015; Mignone et al., 2018; van Marle et al., 2018). We will be concerned with relativistic $e^{\pm}$ plasmas, so PIC methods are very well suited for this problem.

Just like any other plasma code, a PIC code must solve self-consistently Maxwell's equations for the electromagnetic fields $(\boldsymbol{E}, \boldsymbol{B})$ and Newton's equations for the particles. These equations are discretized in time, with a time step Δt. During one time step, three distinct operations must be performed. First the positions and velocities of the particles are

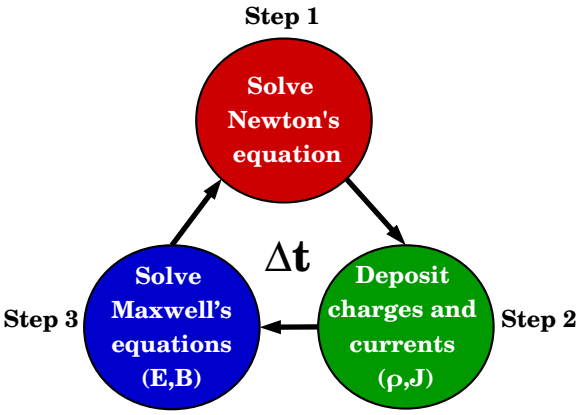


FIGURE 9.1: Typical threefold computation procedure of PIC algorithms per time step. Figure taken from Sironi & Cerutti (2017).

evolved, then the electromagnetic source terms (charge and current densities) are collected on the grid, and finally fields are updated by solving Maxwell's equations on the grid. This procedure is depicted in Fig. 9.1.

9.2 Particle motion

9.2.1 Leapfrog algorithms

In choosing the integration scheme for the particles, there must be a compromise between efficiency (the equations of motion are solved at each time step for every of the N particles in the simulation, with $N \simeq 10^8$), computer storage capacity (high-order integration schemes can store several particles coordinates per time step), and accuracy. A very appropriate choice is the leapfrog scheme (Feynman et al., 1963). Let us illustrate the principle on the simple case of a one-dimensional particle, submitted to a velocity-independent force $F(x)$. Position and velocities are staggered on the time axis by half a timestep $\Delta t/2$: positions are defined at $t_n = n\Delta t$ and velocities/momenta at $t_{n+1/2} = (n+1/2)\Delta t$. At the time step n, the basic integration formula for the leapfrog algorithm is

$$x^{n+1} = x^n + \frac{\Delta t}{m_s}\, p^{n+1/2}, \tag{9.3}$$

$$p^{n+3/2} = p^{n+1/2} + \Delta t F\left(x^{n+1}\right). \tag{9.4}$$

This finite difference scheme is time-centered, so it is time-reversal invariant. As a result, one can prove by Taylor expansion in Δt that this scheme is accurate to second order. Besides, unlike the second-order Runge-Kutta method for instance, it requires only one evaluation of the function F per time step. The initialization of the scheme seems problematic, as for given initial position and momentum (x^0, p^0), one must compute $p^{1/2}$. This induces a first-order error, but this does not matter since it only has to be done once. A minor downside of this scheme is that positions and velocities are never known at the same time step, which might be inconvenient for a variety of diagnostics. This is easily remedied by a minor modification of the leapfrog algorithm, which turns into the "velocity Verlet" algorithm (Verlet, 1967; Swope et al., 1982):

$$p^{n+1/2} = p^n + \frac{\Delta t}{2} F\left(x^n\right), \tag{9.5}$$

$$x^{n+1} = x^n + \frac{\Delta t}{m_s}\, p^{n+1/2}, \tag{9.6}$$

$$p^{n+1} = p^{n+1/2} + \frac{\Delta t}{2} F\left(x^{n+1}\right). \tag{9.7}$$

This scheme is actually equivalent to the leapfrog algorithm, once suitable initial conditions are supplied. Note that the force in Eq. (9.7) needed to compute p^{n+1} is the same as the one needed in Eq. (9.5) to compute $p^{n+3/2}$, so that the Verlet algorithm also necessitates only one computation of F per time step. The principle of the leapfrog is represented in Fig. 9.2.

This is all well and good, but the main upside of leapfrog algorithms is its symplecticity (Young, 2013; Springel, 2016). The time evolution of the phase space coordinates can be seen as a flow in phase space. Let us follow two trajectories initially infinitesimally close, with initial conditions $\left(x^0, p^0\right)$ and $\left(x^0 + \delta x^0, p^0 + \delta p^0\right)$. The coordinate intervals evolve as

$$\begin{pmatrix} \delta x^{n+1} \\ \delta p^{n+1} \end{pmatrix} = J \begin{pmatrix} \delta x^n \\ \delta p^n \end{pmatrix}. \tag{9.8}$$

Eq. (9.8) is actually the discretized version of a coordinate transformation, and J can be seen as the Jacobian of the transformation. An algorithm is said to be symplectic if its Jacobian is symplectic:

$$J^T \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{9.9}$$

where I is the identity matrix. Symplecticity is a property of all Hamiltonian flows (Goldstein et al., 2012). This property implies that the area of a given volume of phase space under the flow of the algorithm is preserved (this is Liouville's theorem). In this simple 1D case, we have $\mathrm{d}V = \mathrm{d}x\,\mathrm{d}p \simeq \delta x^n \delta p^n$ at the time step n. Hence, the algorithm preserves phase space volume if

$$\det J = 1. \tag{9.10}$$

This condition is met by the leapfrog and Verlet algorithms (Young, 2013; Springel, 2016). This implies that they possess global stability. Because phase space conservation is satisfied, there can be no secular trend, and global conserved quantities (energy, angular momentum)

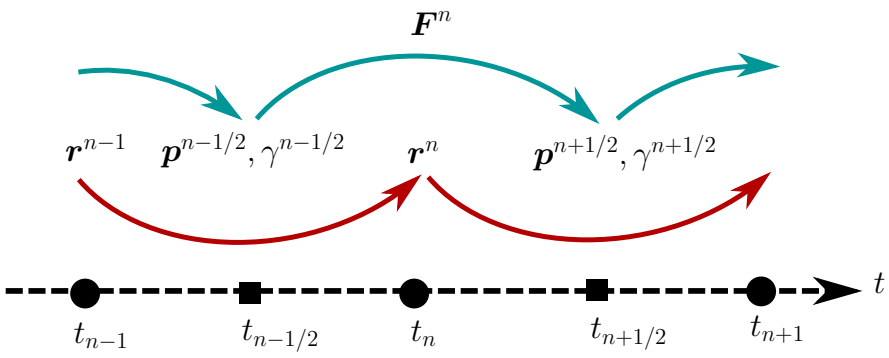


FIGURE 9.2: Illustrative diagram of the leapfrog scheme for the particles. The positions are defined at integer time steps, whereas the velocities are defined at half-integer time steps.

are usually bounded (Hairer et al., 2005). On the contrary, such trends can be oberved on long time scales even in very accurate non-symplectic schemes, such as the fourth order Runge-Kutta scheme.

> The leapfrog and velocity Verlet algorithm that we have presented are accurate to second order. A fourth-order symplectic algorithm has been devised by Forest & Ruth (1990). For the purpose of our simulations, the second-order leapfrog algorithm is sufficient.

9.2.2 Equations of motion

Let us define the particle's 3-velocity $\boldsymbol{v}$ and normalized momentum $\boldsymbol{u} = \gamma\boldsymbol{v}/c = \boldsymbol{p}/m_s c$. The special relativistic versions of Newton's equations read

$$\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t} = \boldsymbol{v} = c\frac{\boldsymbol{u}}{\gamma} \tag{9.11}$$

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} = \frac{q_s}{m_s c}\left(\boldsymbol{E} + \frac{\boldsymbol{u}}{\gamma}\times\boldsymbol{B}\right) + \boldsymbol{g}, \tag{9.12}$$

where $\gamma = \sqrt{1+\boldsymbol{u}^2}$ is the Lorentz factor of the particle, and $\boldsymbol{g}$ is the backreaction radiation force (see Eq. (8.56)), if synchrotron radiation is included in the simulation. Let us momentarily ignore this reaction force. In the spirit of the leapfrog algorithm, these equations are discretized as

$$\frac{\boldsymbol{r}^{n+1} - \boldsymbol{r}^n}{\Delta t} = c\frac{\boldsymbol{u}^{n+1/2}}{\gamma^{n+1/2}}, \tag{9.13}$$

$$\frac{\boldsymbol{u}^{n+1/2} - \boldsymbol{u}^{n-1/2}}{\Delta t} = \frac{q_s}{m_s c}\left(\boldsymbol{E}^n + \frac{1}{2\gamma^n}\left(\boldsymbol{u}^{n+1/2} + \boldsymbol{u}^{n-1/2}\right)\times\boldsymbol{B}^n\right). \tag{9.14}$$

The Lorentz force $\boldsymbol{\mathcal{F}} = q_s\left(\boldsymbol{E} + (\boldsymbol{u}/\gamma)\times\boldsymbol{B}\right)$ must be evaluated at integer time steps $n \in \mathbb{N}$ (see Fig. 9.2), so the mean velocity $\boldsymbol{u}^n \simeq \left(\boldsymbol{u}^{n+1/2} + \boldsymbol{u}^{n-1/2}\right)/2$ is employed. Besides, the fields appearing in these equations are evaluated at the particle's position. A linear interpolation from the grid to the particle's position works fine for this purpose.

Still, one difficulty remains, because the Lorentz force is velocity-dependent. A practical implementation of the particle mover is the Boris push (Boris, 1970; Birdsall & Langdon, 1985), which splits the effects of the electric and magnetic force. With the knowledge of $\boldsymbol{r}_n$ and $\boldsymbol{u}^{n-1/2}$, we define

$$\boldsymbol{u}^- = \boldsymbol{u}^{n-1/2} + \frac{q_s\Delta t}{2cm_s}\boldsymbol{E}^n, \tag{9.15}$$

$$\boldsymbol{u}^+ = \boldsymbol{u}^{n+1/2} - \frac{q_s\Delta t}{2cm_s}\boldsymbol{E}^n, \tag{9.16}$$

which allows us to rewrite the momentum equation into

$$\frac{\boldsymbol{u}^+ - \boldsymbol{u}^-}{\Delta t} = \frac{q_s}{2m_s c\gamma_n}(\boldsymbol{u}^+ + \boldsymbol{u}^-)\times\boldsymbol{B}^n. \tag{9.17}$$

The physical interpretation of Eq. (9.15), (9.16) and (9.17) is the following. In a time-centered fashion, the particle is linearly accelerated by the electric field for half a time step, then has its velocity precessing around the magnetic field by an angle $\simeq \omega_\mathrm{g}\Delta t$ (where $\omega_\mathrm{g} = qB/\gamma m_s c$

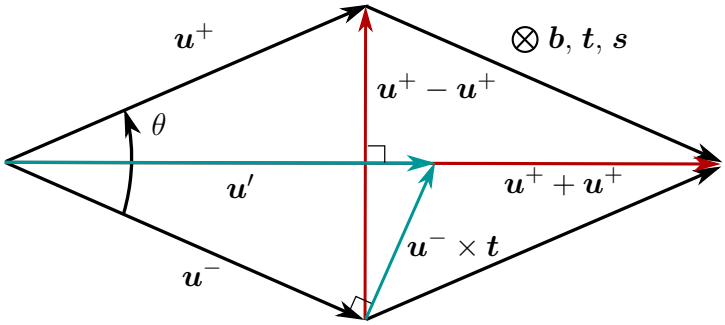


FIGURE 9.3: Geometrical construction of the Boris push. For simplicity, the velocity vectors are assumed to be perpendicular to the magnetic field $\boldsymbol{b}$. The goal is to reconstruct $\boldsymbol{u}^+$, starting from $\boldsymbol{u}^-$. Figure inspired by Birdsall & Langdon (1985).

is the relativistic cyloctron angular frequency), and then reaccelerated for the remaining half time step. All that remains is to get an explicit expression for $\boldsymbol{u}^+$ as a function of $\boldsymbol{u}^-$.

The geometrical construction is represented in Fig. 9.3, in the simple case where $\boldsymbol{B} \perp \boldsymbol{u}$. From Eq. (9.17), we see that the half-angle of rotation is given by

$$t = \tan\frac{\theta}{2} = -\frac{\|\boldsymbol{u}^+ - \boldsymbol{u}^-\|}{\|\boldsymbol{u}^+ + \boldsymbol{u}^-\|} = -\frac{q_s B \Delta t}{2 m_s c \gamma_n} = -\frac{\omega_g \Delta t}{2}, \tag{9.18}$$

where $\gamma_n = \sqrt{1 + \boldsymbol{u}_-^2}$. This matches the physical rotation angle if $\omega_g \Delta t \ll 1$ only, indicating that the cyclotron frequency must be time-resolved by the simulation to follow accurately the motion of particles. Finally, we define the unit vector $\boldsymbol{b} = \boldsymbol{B}/B$ and the auxiliary vectors $\boldsymbol{t} = -\tan(\theta/2)\boldsymbol{b} = -t\boldsymbol{b}$ and $\boldsymbol{s} = 2\boldsymbol{t}/\left(1 + \boldsymbol{t}^2\right)$. Using the tangent half-angle formulas, we have $\sin\theta = 2t/\left(1 + t^2\right)$ and $\cos\theta = \left(1 - t^2\right)/\left(1 + t^2\right)$. Then, one can check that the vector

$$\boldsymbol{u}' = \boldsymbol{u}^- + \boldsymbol{u}^- \times \boldsymbol{t}, \tag{9.19}$$

is orthogonal to $\boldsymbol{u}^+ - \boldsymbol{u}^-$, and consequently, parallel to $\boldsymbol{u}^+ + \boldsymbol{u}^-$. Since $\boldsymbol{u}^+ - \boldsymbol{u}^-$ is parallel to $\boldsymbol{u}' \times \boldsymbol{B}$, $\boldsymbol{u}^+$ can eventually be expressed as

$$\boldsymbol{u}^+ = \boldsymbol{u}^- + \boldsymbol{u}' \times \boldsymbol{s}, \tag{9.20}$$

which fulfills the requirement that $\|\boldsymbol{u}^+\| = \|\boldsymbol{u}^-\|$. All in all $\boldsymbol{u}^{n+1/2}$ can be expressed as a function of $\boldsymbol{u}^{n-1/2}$ in the leapfrog scheme.

Other Lorentz force pushers have been implemented; a comprehensive review can be found in Ripperda et al. (2018). In particular, Ripperda et al. (2018) show that the Boris push only conserves the total numerical energy in the absence of an electric field. Actually, the Boris push is not symplectic. Yet, it performs outstandingly even in the presence of an electric field, because it still preserves phase space volume (Qin et al., 2013). This brings long-term stability to the Boris push.

The radiation-reaction force is included using the scheme detailed in Tamburini et al. (2010). The full leapfrog time step is

$$\frac{\boldsymbol{u}^{n+1/2} - \boldsymbol{u}^{n-1/2}}{\Delta t} = \boldsymbol{F}_{\mathrm{L}}^n + \boldsymbol{g}^n. \tag{9.21}$$

The scheme consists in pushing the momentum of the particle independently with the Lorentz force $\boldsymbol{F}_{\mathrm{L}}$ and the reaction force $\boldsymbol{g}$, with

$$\frac{\boldsymbol{u}_{\mathrm{L}}^{n+1/2}-\boldsymbol{u}^{n-1/2}}{\Delta t}=\frac{1}{m_s c}\boldsymbol{F}_{\mathrm{L}}^{n};\quad \frac{\boldsymbol{u}_{\mathrm{R}}^{n+1/2}-\boldsymbol{u}^{n-1/2}}{\Delta t}=\frac{1}{m_s c}\boldsymbol{g}_n. \tag{9.22}$$

and then to reassemble the full updated momentum $\boldsymbol{u}^{n+1/2}=\boldsymbol{u}_{\mathrm{L}}^{n+1/2}+\boldsymbol{u}_{\mathrm{R}}^{n+1/2}-\boldsymbol{u}^{n-1/2}=\boldsymbol{u}_{\mathrm{L}}^{n+1/2}+\boldsymbol{g}^n\Delta t/m_s c$. The Boris push can still bu used to evolve the momentum under the action of the Lorentz force. In order to estimate $\boldsymbol{g}^n$, the Lorentz factor of the particle at an integer time step is needed. The total momentum at $t_n=n\Delta t$ is estimated as $\boldsymbol{u}^n\simeq(\boldsymbol{u}_{\mathrm{L}}^{n+1/2}+\boldsymbol{u}^{n-1/2})/2$.

9.2.3 Source terms

In order to solve Maxwell's equations, the source terms are yet to be computed. The charge density and current density are determined by the particle contributions:

$$\rho(\boldsymbol{r})=\sum_{k=1}^{N}q_k w_k\delta(\boldsymbol{r}-\boldsymbol{r}_k), \tag{9.23}$$

$$\boldsymbol{j}(\boldsymbol{r})=\sum_{k=1}^{N}q_k w_k\boldsymbol{v}_k\delta(\boldsymbol{r}-\boldsymbol{r}_k). \tag{9.24}$$

Then, the values of the source terms must be deposited on the grid nodes, so they can be injected on the spatially discretized Maxwell's equations. There exist several interpolation schemes, but a linear one is sufficient most of the time. In 2D Cartesian coordinates (x,y), if a macroparticle with charge q and weight w is located between the nodes $x\in[x_i,x_{i+1}]$ and $y\in[y_j,y_{j+1}]$, then the charge that is deposited on each node is

$$q_{i,j}=(1-f_x)(1-f_y)\,q\,w,\quad q_{i,j+1}=(1-f_x)f_y\,q\,w, \tag{9.25}$$

$$q_{i+1,j}=f_x(1-f_y)\,q\,w,\qquad q_{i+1,j+1}=f_x f_y\,q\,w, \tag{9.26}$$

where f_x and f_y are defined as

$$f_x=\frac{x-x_i}{x_{i+1}-x_i},\ f_y=\frac{y-y_j}{y_{j+1}-y_j}. \tag{9.27}$$

This has been dubbed an "area weighting" procedure, since the weights f_x and f_y are given by the ratio of the surfaces linearly defined by the position of the particle within the cell (see Fig. 9.4). It gives the particles an effective size, which amounts to replacing the Dirac function in Eq. (9.23) and (9.24) by a wider shape function which smooths the source terms. For example, a first-order interpolation scheme involves a triangular shape function. Due to this procedure, scales smaller than the grid resolution cannot be resolved. The same linear interpolation is used in the Boris push to interpolate the fields from the grid to the location of the particles.

9.3 Field evolution

Now that the source terms have been updated, the last step consists in evolving the fields using Ampère's and Faraday's equations:

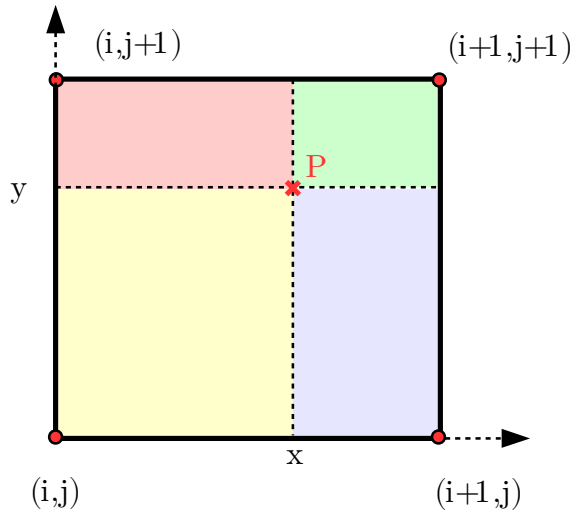


FIGURE 9.4: Area-weighting procedure on the case of a 2D Cartesian grid. The fractional charge of a particle at P is deposited on the four surrounding nodes in proportion to the colored shaded areas. Figure adapted from Sironi & Cerutti (2017).

$$\frac{\partial \boldsymbol{E}}{\partial t} = c\boldsymbol{\nabla} \times \boldsymbol{B} - 4\pi \boldsymbol{J} \tag{9.28}$$

$$\frac{\partial \boldsymbol{B}}{\partial t} = -c\boldsymbol{\nabla} \times \boldsymbol{E}. \tag{9.29}$$

The most common approach in PIC codes is the finite difference time domain method (Yee, 1966). Let us illustrate this method in the case of a 2D Cartesian grid (x, y), assuming invariance along the third z axis.

9.3.1 Time discretization

Like the particle pushers, a time-centered leapfrog method is preferable for the time evolution of the fields. The joint evolution of the particles and fields is shown in Fig. 9.5. In order to advance the particles, the electric and magnetic fields must be known at the same time. This is precisely the case where the velocity Verlet algorithm comes in handy, since it allows us to know both $\boldsymbol{E}$ and $\boldsymbol{B}$ at all integer time steps (taking the momentum for the magnetic field and the position for the electric field). The magnetic field is first advanced only by half a time step $\Delta t/2$, so the particle motion can be computed. Then, the magnetic field is once again advanced by $\Delta t/2$, completing the Verlet algorithm. Assuming $\boldsymbol{E}^n$ and $\boldsymbol{B}^n$ are initially known, the various steps read

1. $\boldsymbol{u}^{n-1/2} \rightarrow \boldsymbol{u}^{n+1/2}$ with $\boldsymbol{E}^n$ and $\boldsymbol{B}^n$.
2. $\boldsymbol{r}^n \rightarrow \boldsymbol{r}^{n+1}$ with $\boldsymbol{u}^{n+1/2}$.
3. $\boldsymbol{B}^{n+1/2} = \boldsymbol{B}^n - \dfrac{c\Delta t}{2} (\boldsymbol{\nabla} \times \boldsymbol{E})^n$.
4. $\boldsymbol{E}^{n+1} = \boldsymbol{E}^n + c\Delta t\, (\boldsymbol{\nabla} \times \boldsymbol{B})^{n+1/2} - 4\pi \Delta t \boldsymbol{J}^{n+1/2}$.
5. $\boldsymbol{B}^{n+1} = \boldsymbol{B}^{n+1/2} - \dfrac{c\Delta t}{2} (\boldsymbol{\nabla} \times \boldsymbol{E})^{n+1}$.

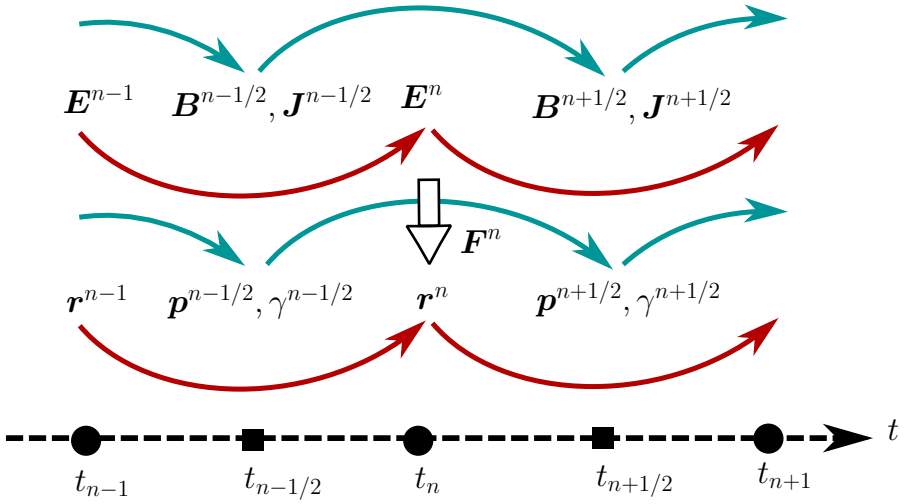


FIGURE 9.5: Simultaneous leapfrogs for the electromagnetic fields and the particles. The Lorentz force $\boldsymbol{F}$ is computed from the fields to push the particles.

The only overlooked subtlety here has been the knowledge of the current density at the time step $n+1/2$. $\boldsymbol{J}^{n+1/2}$ is computed as the average betwwen $\boldsymbol{J}\left(\boldsymbol{r}^{n}, \boldsymbol{u}^{n+1/2}\right)$ (before the coordinates of the particles are updated) and $\boldsymbol{J}\left(\boldsymbol{r}^{n+1}, \boldsymbol{u}^{n+1/2}\right)$ (after the coordinates are updated).

9.3.2 Spatial discretization

The curl and divergence of the electromagnetic fields are estimated by a centered finite-difference method, so this method has second-order accuracy. The key idea is to stagger the different components of the fields on different nodes of the grid, as pictured in Fig. 9.6 for a 2D Cartesian grid, with regular spacing Δx and Δy. This spatial staggering is often dubbed the "Yee grid". For example, the components of Eq. (9.29) discretized in this fashion are

$$\frac{(B_x)^{n+1/2}_{i,j+1/2} - (B_z)^{n-1/2}_{i,j+1/2}}{\Delta t} = -c\frac{(E_z)^{n}_{i,j+1} - (E_z)^{n}_{i,j}}{\Delta y}, \tag{9.30}$$

$$\frac{(B_y)^{n+1/2}_{i+1/2,j} - (B_y)^{n-1/2}_{i+1/2,j}}{\Delta t} = c\frac{(E_z)^{n}_{i+1,j} - (E_z)^{n}_{i,j}}{\Delta x}, \tag{9.31}$$

$$\begin{aligned}\frac{(B_z)^{n+1/2}_{i+1/2,j+1/2} - (B_z)^{n-1/2}_{i+1/2,j+1/2}}{\Delta t} &= -c\frac{(E_y)^{n}_{i+1,j+1/2} - (E_y)^{n}_{i,j+1/2}}{\Delta x} \\ &\quad + c\frac{(E_x)^{n}_{i+1/2,j+1} - (E_z)^{n}_{i+1/2,j}}{\Delta y},\end{aligned} \tag{9.32}$$

where the upper index denotes the time step and the lower index denotes the spatial position on the Yee grid.

What about the two other Maxwell's equations? One of the major benefit of employing the Yee grid is that the equation $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$ is automatically satisfied to machine roundoff precision, provided the simulation is initialized with a divergence-free magnetic field.

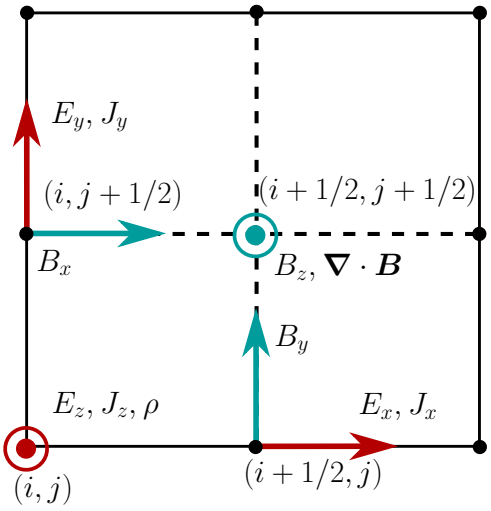


FIGURE 9.6: 2D Cartesian Yee grid. The red arrows denote the components of the electric fields, the blue ones denote the magnetic fields. The positions on the mesh are labelled by the indices i in the x direction and j in the y direction.

Proof

The discretized divergence of $\boldsymbol{B}$ is given by

$$(\boldsymbol{\nabla}\cdot\boldsymbol{B})^{n+1/2}_{i+1/2,j+1/2} = \frac{(B_x)^{n+1/2}_{i+1,j+1/2} - (B_x)^{n+1/2}_{i,j+1/2}}{\Delta x} + \frac{(B_y)^{n+1/2}_{i+1/2,j+1} - (B_y)^{n+1/2}_{i+1/2,j}}{\Delta y}. \quad (9.33)$$

Let us compute the variation $\delta\left(\boldsymbol{\nabla}\cdot\boldsymbol{B}\right) = (\boldsymbol{\nabla}\cdot\boldsymbol{B})^{n+1/2}_{i+1/2,j+1/2} - (\boldsymbol{\nabla}\cdot\boldsymbol{B})^{n-1/2}_{i+1/2,j+1/2}$, using the discretized equations (9.30), (9.31) and (9.32):

$$\begin{aligned}
\delta\left(\boldsymbol{\nabla}\cdot\boldsymbol{B}\right) &= \frac{(B_x)^{n+1/2}_{i+1,j+1/2} - (B_x)^{n-1/2}_{i+1,j+1/2}}{\Delta x} - \frac{(B_x)^{n+1/2}_{i,j+1/2} - (B_x)^{n-1/2}_{i,j+1/2}}{\Delta x} \\
&+ \frac{(B_y)^{n+1/2}_{i+1/2,j+1} - (B_y)^{n-1/2}_{i+1/2,j+1}}{\Delta y} - \frac{(B_y)^{n+1/2}_{i+1/2,j} - (B_y)^{n-1/2}_{i+1/2,j}}{\Delta y} \quad (9.34)\\
&= \frac{c\Delta t}{\Delta x \Delta y}\big(- (E_z)^n_{i+1,j+1} + (E_z)^n_{i+1,j} + (E_z)^n_{i,j+1} - (E_z)^n_{i,j} \\
&+ (E_z)^n_{i+1,j+1} - (E_z)^n_{i,j+1} - (E_z)^n_{i+1,j} + (E_z)^n_{i,j}\big) \quad (9.35)\\
&= 0. \quad (9.36)
\end{aligned}$$

As a result, the divergence of $\boldsymbol{B}$ is constant. It remains zero if it is initialized at zero. This result holds only in a spatially and time centered scheme.

Unfortunately, this scheme does not enforce Maxwell-Gauss' equation $\boldsymbol{\nabla}\cdot\boldsymbol{E} = 4\pi\rho$, and the electric charge deposited *on the grid* is not exactly conserved. If this is left untended, small errors can accumulate and lead to unphysical charge densities. Therefore, the electric field has to be periodically corrected (Birdsall & Langdon, 1985; Hockney & Eastwood, 1988), a method called "divergence cleaning". Let $\boldsymbol{E}_0$ be the electric field provided by the field solver, and ρ the charge density. The corrected electric field is $\boldsymbol{E} = \boldsymbol{E}_0 - \boldsymbol{\nabla}\phi$, where ϕ obeys

Poisson's equation $\Delta\phi = -4\pi\rho + \boldsymbol{\nabla}\cdot\boldsymbol{E}_0$. The electric potential is defined at the nodes (i,j). Denoting $S_{i,j}$ the source term on the right-hand side, the discretized form of Poisson's equation is

$$\frac{\phi_{i+1,j}+\phi_{i-1,j}-2\phi_{i,j}}{\Delta x^2}+\frac{\phi_{i,j+1}+\phi_{i,j-1}-2\phi_{i,j}}{\Delta y^2}=S_{i,j}, \tag{9.37}$$

which can be recast into

$$\phi_{i,j}=\frac{\Delta y^2\left(\phi_{i+1,j}+\phi_{i-1,j}\right)+\Delta x^2\left(\phi_{i,j+1}+\phi_{i,j-1}\right)}{2\left(\Delta x^2+\Delta y^2\right)}-\frac{\Delta x^2\Delta y^2}{2\left(\Delta x^2+\Delta y^2\right)}S_{i,j}. \tag{9.38}$$

This equation yields an implicit value for $\phi_{i,j}$. To solve it, a computationally easy way is to use an iterative method. The most basic one is the Jacobi method (Press et al., 1992). After an initial guess $\phi_{i,j}^{[0]}$, an updated potential is obtained by applying Eq. (9.38), so that the estimate of $\phi_{i,j}$ after k iterations is

$$\phi_{i,j}^{[k]}=\frac{\Delta y^2\left(\phi_{i+1,j}^{[k-1]}+\phi_{i-1,j}^{[k-1]}\right)+\Delta x^2\left(\phi_{i,j+1}^{[k-1]}+\phi_{i,j-1}^{[k-1]}\right)}{2\left(\Delta x^2+\Delta y^2\right)}-\frac{\Delta x^2\Delta y^2}{2\left(\Delta x^2+\Delta y^2\right)}S_{i,j}. \tag{9.39}$$

It can be shown that for this particular problem, the procedure converges for any initial guess (LeVeque, 2007). A number of iterations between 5 and 10 is usually enough. This method has the benefit of allowing generalizations to more involved geometries (cylindrical or spherical for example). There exist some schemes to reconstruct charge-conserving ρ and $\boldsymbol{J}$ (Villasenor & Buneman, 1992; Esirkepov, 2001), but they are limited to Cartesian geometries for now.

9.3.3 Grid resolution

A question that remains to be adressed is the choice of Δt and the grid resolution. A general method to assess the stability of a numerical sheme consists in evaluating the behaviour of plane waves (Birdsall & Langdon, 1985). Injecting the fields $(\boldsymbol{E},\boldsymbol{B})(t,\boldsymbol{r}) = (\boldsymbol{E}_0,\boldsymbol{B}_0)\exp(i(\omega t-\boldsymbol{k}\cdot\boldsymbol{r}))$ into the discretized versions of Eqs. (9.28) and (13.10), one finds:

$$\left(\frac{\sin(\omega\Delta t/2)}{c\Delta t}\right)^2=\left(\frac{\sin(k_x\Delta x/2)}{\Delta x}\right)^2+\left(\frac{\sin(k_y\Delta y/2)}{\Delta y}\right)^2. \tag{9.40}$$

More details on this derivation are given in the general case in Sec. 10.2.4. This dispersion relation determines the frequency ω of the wave as a function of the wave vector. It reduces to the light wave dispersion relation $\omega = kc$ for $\Delta t \to 0$. The scheme is stable if and only if ω is real, which yields the *Courant-Friedrichs-Levy* (CFL) condition (Pritchett, 2003; Sironi & Cerutti, 2017):

$$\mathcal{C}=(c\Delta t)^2\left(\frac{1}{\Delta x^2}+\frac{1}{\Delta y^2}\right)<1, \tag{9.41}$$

with $\mathcal{C}$ the CFL number. Usually, the grid resolution is given by the user, and Eq. (9.41) is used to prescribe the time step *a posteriori* with a CFL number lower than 1. The most critical necessity is that the Debye length, which is just the skin depth d_e in relativistic plasmas, should be resolved. For a plasma of typical density n_0, the skin depth reads

$$d_e=\sqrt{\frac{m_s c^2}{4\pi n_0 {q_s}^2}}. \tag{9.42}$$

Spatially resolving the skin depth is equivalent to resolving the plasma frequency in the time domain. Failure to do so results in the appearance of spurious electric fields, which can accelerate particles and heat the plasma. Basically, the initial assumption of the PIC method that all physics below the Debye scale could be neglected breaks down. The Larmor radius of particles should also be resolved, although the consequences are less dramatic if it is not.

9.4 Parallelization

The most common parallelization strategy is to use a grid-based domain decomposition, applied through the Message Passing Interface (MPI), which ensures the communication between the different cores. The simulation box is divided into smaller domains, following a Cartesian topology. Each of the N CPUs is assigned to a single domain, its task being to evolve the fields on the sub-grid, and to push the particles propagating within the domain. Each CPU goes through all the previously described steps of the PIC algorithm. At the end of each step, it must communicate with its neighbours several pieces of information: which particles have left its domain, which particles are entering the domain, and what are the updated values of the fields at the boundaries of the domain. In a Cartesian 2D decomposition, the bulk domains have 8 neighbours.

Communication of the grid values is performed through the intermediary of "ghost cells", which are organized as shown in Fig. 9.7 in a 1D geometry. Let us consider a domain which is not in contact with the boundary of the global simulation box, with its nodes labeled from 1 to N (Proc. 2 in the figure). The nodes 1 and N are out of the physical domain. Loops in spatial indices (in the field pusher for example) run from 2 to $N-1$. To compute the finite-difference expressions at the nodes 2 and $N-1$, the values of the fields are needed at 1 and N. These values are provided by the neighboring domains (Proc. 1 and 3 in the figure): the domains overlap in such a way that they share a ghost cell. In the figure, at the end of the loops, the node 1 of Proc. 2 is updated by the updated value at the node $N-1$ of Proc. 1, while the node N of Proc. 1 is updated using the updated value at the node 2 of Proc. 2.

PIC simulations are very easily scalable to a large number of cores: aside from what happens at the boundaries of the domains, all operations can be performed indepently. This makes them useful for astrophysical simulations, which are plagued by a catastrophically large separation of scales, from the microscopic kinetic scales to the largest scales associated with the global system. In a typical run, most of the computation time is spent pushing the particles. As a result, load balancing issues can occur if a small number of CPUs have to take care of the majority of the particles. This is a major downside of grid-based decompositions.

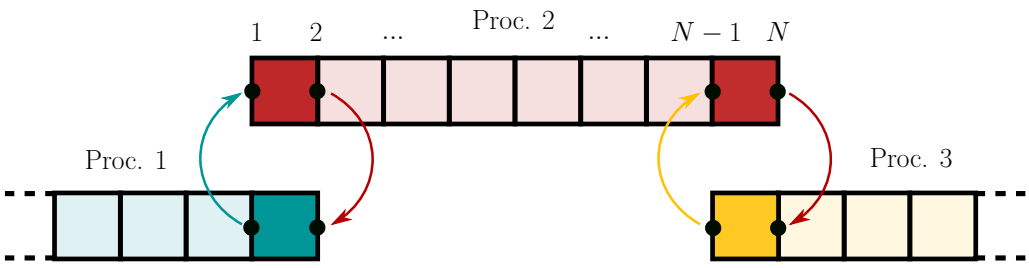


FIGURE 9.7: Diagram illustrating the communication between neighboring 1D domains. The fields are defined at the edges, with N nodes per domain. As an example, the blue arrow indicates that the field value at node 1 in Proc. 2 is assigned the value from node $N-1$ in Proc. 1. The darker cells are the "ghost cells".

For this parallelization technique to work efficiently, particles have to be distributed quite uniformly across all domains. This must be taken into account when designing the geometry of the domain decomposition.

> Some PIC codes, such as `Smilei`, use involved techniques to avoid load balancing issues (Derouillat et al., 2018). Examples are dynamic load balancing, or the use of the OpenMP library to benefit from the shared memory within one node.

Chapter 10

The GRZeltron upgrade

We have seen in Sec. 9 the basics of PIC simulations. The first PIC code which is both multidimensional and capable of fully handling general relativity was written by Kyle Parfrey, Sasha Philippov and Benoît Cerutti. General relativity is encoded in the $3+1$ formalism, which has allowed them to rely heavily on the architecture of the previously existing special-relativistic version of `Zeltron`, one of the few PIC codes designed for non-Cartesian geometries. `GRZeltron` was first introduced in Parfrey et al. (2019), although no description of the underlying numerical methods has been published so far. In this section, I present some of the core ideas and numerical methods used in `GRZeltron`, which is the code that I have used throughout my thesis.

The code works in spherical KS coordinates (t, r, θ, φ), because these encode the spherical symmetry of the problem and remain regular at the event horizon, unlike spherical BL coordinates. All quantities in the code are invariant by rotation around the spin axis of the black hole. In that sense, although particles move in a 3D space, the fields are defined by 2D arrays. Such axisymmetric simulations are often called "2.5D".

10.1 Particle solver

10.1.1 Equations of motion

We need to derive a geodesic equation for charged massive particles which can fit with the $3+1$ formulation used in the code. The objective is to take advantage of the methods already implemented in special relativistic PIC codes, so we seek first-order equations as similar as possible to Eq. (9.11) and (9.12). With this in mind, a sensible idea is to use a Hamiltonian formalism and massage Hamilton's equations into a tractable form: this is done in the Appendix 10.A. The equations of motion ensuing from this procedure for particles of mass m_e and charge q are:

$$\frac{1}{c}\frac{\mathrm{d}x^i}{\mathrm{d}t} = \frac{\alpha}{\Gamma}h^{ij}u_j - \beta^i \tag{10.1}$$

$$\frac{1}{c}\frac{\mathrm{d}u_i}{\mathrm{d}t} = -\Gamma\partial_i\alpha + u_j\partial_i\beta^j - \frac{\alpha}{2\Gamma}u_j u_k \partial_i h^{jk} + \frac{q}{m_\mathrm{e}c^2}\alpha\left(h_{ij}D^j + e_{ijk}\frac{h^{jl}u_l}{\Gamma}B^k\right). \tag{10.2}$$

These equations have been derived by direct manipulations of the geodesic equation (see Eq. (10.47)) in the case of massless particles by Hughes et al. (1994). Dodin & Fisch (2010) derived the equations for charged particles in a quite different form. The Hamiltonian formalism was used by Bacchini et al. (2018, 2019) to design a symplectic and energy-conserving particle pusher. The last term on the right-hand side of Eq. (10.2) is the Lorentz force $\boldsymbol{\mathcal{F}}$ exerted on the particles, written in terms of the fields measured by FIDOs.

The first three terms on the right-hand side of Eq. (10.2) are of gravitational origin, arising from to the curvature of spacetime. In the weak-field limit, the linearized version of this equation supports a "gravitomagnetic" interpretation: α behaves as a scalar potential and

$\boldsymbol{\beta}$ as a vector potential. In this limit, α is analog to the gravitational potential, whereas $\boldsymbol{\beta}$ causes a deflection of particle trajectories, the smoking gun of frame dragging.

Eq. (10.1) highlights the difference between the physical velocity of the particle, measured by the FIDO, and the velocity as measured on the grid (see Vincent et al. 2012 for a discussion). By definition, the 3-velocity $\boldsymbol{v}$ in the global coordinate basis has components

$$v^i = \frac{\mathrm{d}x^i}{\mathrm{d}t}; \tag{10.3}$$

it is the rate of change of the coordinates of the particle with respect to the grid. The 3-velocity $\boldsymbol{V}$ measured by the FIDO is

$$V^i = \frac{\mathrm{d}l^i}{\mathrm{d}\tau_{\mathrm{FIDO}}}, \tag{10.4}$$

where l^i is the displacement vector of the particle with respect to the FIDO, and $\mathrm{d}\tau_{\mathrm{FIDO}}$ is the increment of the FIDO's proper time between t and $t + \mathrm{d}t$. The situation is represented in Fig. 10.1. Now, if a 4-vector X^μ is projected onto a hypersurface with the orthogonal projector $h^\mu{}_\nu$ (see Eq. (5.1)), its components become $\overline{X}^\mu = h^\mu{}_\nu X^\nu = \left(0, X^i + \beta^i X^0\right)$. From the vantage point of the FIDO, the origin of the coordinate system is shifted by an amount $c\beta^i\,\mathrm{d}t$. Consequently, the displacement vector measured by the FIDO between t and $t + \mathrm{d}t$ is $\mathrm{d}l^i = \mathrm{d}x^i + \beta^i c\,\mathrm{d}t$. Using the relation between the FIDO's proper time and the universal time $\mathrm{d}\tau_{\mathrm{FIDO}} = \alpha\,\mathrm{d}t$ (Eq. (5.12)), we find

$$V^i = \frac{1}{\alpha}\left(v^i + c\beta^i\right) \iff v^i = \alpha V^i - c\beta^i. \tag{10.5}$$

We notice that this is precisely the first equation of motion (10.1), with $V^i = h^{ij}u_j c/\Gamma$. In particular, the Lorentz force can be written more compactly as

$$\boldsymbol{\mathcal{F}} = q\left(\boldsymbol{D} + \boldsymbol{V} \times \boldsymbol{B}\right), \tag{10.6}$$

completing the reformulation of the equations of motion into a form similar to the ones in

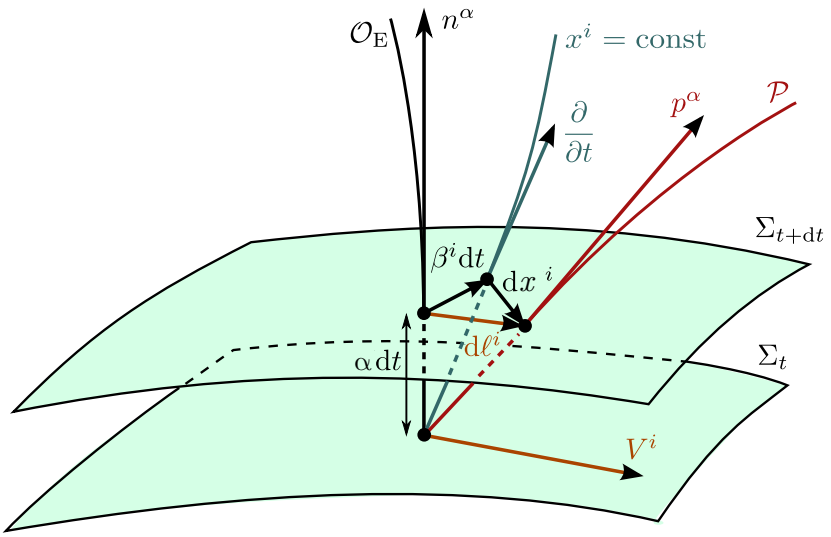


FIGURE 10.1: Diagram illustrating the construction of the 3-velocity measured by a FIDO V^i, and its relation to the coordinate-basis velocity $v^i = \mathrm{d}x^i/\mathrm{d}t$. $\mathcal{P}$ is the wordline of the particle, and $\mathcal{O}_{\mathrm{E}}$ is that of the FIDO. Figure adapted from Vincent et al. (2012).

flat spacetime. This is the usual expression of the Lorentz force, where the electric field, magnetic field and velocity are all measured by the FIDO. Eq. (10.5) also implies that the electric current density measured by the FIDO $\boldsymbol{J}$ is related to the one measured in the coordinate basis $\boldsymbol{j}$ by $\boldsymbol{j} = \alpha\boldsymbol{J} - c\rho\boldsymbol{\beta}$, as mentioned earlier (see Eq. (5.50)).

10.1.2 Splitting

Unlike the special relativistic case (Eq. (9.12)), the evolution equation of the momentum (10.2) involves two very different terms : the gravitational force operator $\mathcal{O}^{\mathrm{g}}$ and the Lorentz force operator $\mathcal{O}^{\mathrm{em}}$, which both depend on the momentum itself. Instead of straightforwardly solving the full system, the philosophy of our code is to retain most of the Zeltron architecture, in particular the excellent features of the Boris pusher. To do so, we split the system into two evolution operators (Hairer et al., 2005). The general evolution of the system $U = [u_1, u_2, u_3]$ is written as

$$U\left(t + \Delta t\right) = \mathcal{O}^{\mathrm{tot}}_{\Delta t}\left[U\left(t\right)\right] \tag{10.7}$$

The crudest idea to solve the joint evolution would be to solve first for $\mathcal{O}_{\mathrm{g}}$ during Δt, then for $\mathcal{O}_{\mathrm{em}}$ during Δt, so that the full evolution would read

$$U\left(t + \Delta t\right) = \left(\mathcal{O}^{\mathrm{em}}_{\Delta t} \circ \mathcal{O}^{\mathrm{g}}_{\Delta t}\right)\left[U\left(t\right)\right]. \tag{10.8}$$

This is called Lie-Trotter splitting (Trotter, 1959). This is only a first-order scheme, and the error with respect to the true evolution is $O\left(\Delta t^2\right)$ (Springel, 2016). This poor performance was readily hinted by the lack of symmetry between $\mathcal{O}_{\mathrm{g}}$ and $\mathcal{O}_{\mathrm{em}}$ in Eq. (10.8), whereas Eq. (10.7) is symmetric. An improvement of this scheme is the Strang splitting (Strang, 1968), which is a symmetrized version of Eq. (10.8):

$$U\left(t + \Delta t\right) = \left(\mathcal{O}^{\mathrm{em}}_{\Delta t/2} \circ \mathcal{O}^{\mathrm{g}}_{\Delta t} \circ \mathcal{O}^{\mathrm{em}}_{\Delta t/2}\right)\left[U\left(t\right)\right]. \tag{10.9}$$

The Strang splitting has an error $O\left(\Delta t^3\right)$ with respect to the true evolution (Springel, 2016).

Example

This splitting is more easily illustrated in the case where the two operators are linear: let us call A the matrix of the gravitational operator and B the matrix of the electromagnetic operator, in the basis $[u_1, u_2, u_3]$. The exact evolution of the system is

$$\dot{U} = AU + BU \Rightarrow U\left(t + \Delta t\right) = \mathcal{O}^{\mathrm{tot}}_{\Delta t} U\left(t\right) = \exp\left(\left(A + B\right)\Delta t\right) U\left(t\right). \tag{10.10}$$

The Lie-Trotter splitting consists in approximating $U_1\left(t + \Delta t\right)$ by $\exp\left(A\Delta t\right)\exp\left(B\Delta t\right)U\left(t\right)$, whereas the Strang splitting defines the evolved vector as $U_2\left(t + \Delta t\right) \simeq \exp\left(A\Delta t/2\right)\exp\left(\Delta t\, B\right)\exp\left(A\Delta t/2\right)U\left(t\right)$. The relative error in the the evolved vector in the case of the Lie-Trotter splitting is

$$\frac{U_2\left(t + \Delta t\right) - U_1\left(t + \Delta t\right)}{U\left(t\right)} = \exp\left(A\Delta t\right)\exp\left(B\Delta t\right) - \exp\left(\left(A + B\right)\Delta t\right) \tag{10.11}$$

$$\simeq \frac{1}{2}\Delta t^2\left(AB - BA\right). \tag{10.12}$$

In the Strang splitting, the second-order term cancels (this can be seen using the Baker-Campbell-Haussdorf formula expanding $\exp\left(\left(A + B\right)\Delta t\right)$).

Note also that the Verlet algorithm (Eq. (9.5), (9.6), (9.7)) was already a special case of Strang splitting, where the vector U was (x, p), with the evolution of x and p being split.

10.1.3 Time evolution

The Boris pusher solves the electromagnetic part $\mathcal{O}^{\mathrm{em}}$. The coordinate push is done simultaneously with the momentum push under the action of the gravitational operator $\mathcal{O}^{\mathrm{g}}$. Because the two equations of motion are not separable, the standard leapfrog method cannot be used: $\mathrm{d}x^i/\,\mathrm{d}t$ depends on x^i and u_i both for example. Instead, we use the implicit midpoint method, which is also second-order accurate. We have a system of ordinary differential equations $\dot{U} = f(U)$ governing the evolution of the vector $U = [x^1, x^2, x^3, u_1, u_2, u_3]$. The implicit midpoint method reads

$$U^{n+1} = U^n + \Delta t f\left(U^{n+1/2}\right), \tag{10.13}$$

where $U^{n+1/2}$ is estimated as

$$U^{n+1/2} = \frac{1}{2}\left(U^n + U^{n+1}\right). \tag{10.14}$$

This is an implicit method, because U^{n+1} is not known at the time of the calculation. Eq. (10.13) is solved iteratively, with the estimation of ${U^{n+1}}^{[k]}$ being

$$ {U^{n+1}}^{[k]} = U^n + \Delta t f\left(\frac{1}{2}\left(U^n + {U^{n+1}}^{[k-1]}\right)\right). \tag{10.15}$$

The method generally converges well after 10 iterations. The derivatives of the coefficients of the metric at x^i are computed numerically by finite differences centered on x^i. All in all, combined with the Strang splitting of the equations of motion, the time evolution of a particle with coordinates $\left({x^i}^n, {u_i}^n\right)$ at the time step n reads as follows.

1. ${u_i}^n \xrightarrow{\mathcal{O}^{\mathrm{em}}_{\Delta t/2}} {u_{i1}}^{n+1/2}$ at position ${x^i}^n$, using the Boris push.
2. ${x^i}^n \to {x^i}^{n+1}$ and ${u_{i1}}^{n+1/2} \xrightarrow{\mathcal{O}^{\mathrm{g}}_{\Delta t}} {u_{i2}}^{n+1/2}$ using the implicit midpoint method.
3. ${u_{i2}}^{n+1/2} \xrightarrow{\mathcal{O}^{\mathrm{em}}_{\Delta t/2}} {u_i}^{n+1}$ at position ${x^i}^{n+1}$, using the Boris push.

10.1.4 Tetrads

The Boris pusher was designed to work with Cartesian coordinates in flat spacetime. However, all tensor quantities in the code are expressed in the spherical KS coordinate basis. In order to preserve the Boris pusher, it is necessary to go from the coordinate basis to a Cartesian basis. To do so, we will use the concept of tetrads (Blau, 2018). From the natural coordinate basis $\mathrm{d}x^\mu$ of 1-forms, a general change of basis reads

$$e^{\hat{a}} = {e^{\hat{a}}}_{\mu}\,\mathrm{d}x^\mu \iff \mathrm{d}x^\mu = {e^\mu}_{\hat{a}} e^{\hat{a}}, \tag{10.16}$$

where ${e^{\hat{a}}}_{\mu}$ is invertible, and ${e^\mu}_{\hat{a}}$ is the inverse matrix of ${e^{\hat{a}}}_{\mu}$ (${e^{\hat{a}}}_{\mu}{e^\mu}_{\hat{b}} = {\delta^{\hat{a}}}_{\hat{b}}$, ${e^\mu}_{\hat{a}}{e^{\hat{a}}}_{\nu} = {\delta^\mu}_\nu$). The matrix ${e^{\hat{a}}}_{\mu}$ is called a tetrad, and ${e^\mu}_{\hat{a}}$ is the inverse tetrad. Importantly, the new base $\left(e^{\hat{a}}\right)$ is *not* a coordinate basis in general. If it were, one could write ${e^{\hat{a}}}_{\mu} = \partial y^{\hat{a}}/\partial x^\mu$ for some coordinate system $y^{\hat{a}}$, which would imply $\partial_\nu {e^{\hat{a}}}_{\mu} = \partial_\mu {e^{\hat{a}}}_{\nu}$: this is not true in general. This operation is a change of basis, not a coordinate transformation. Similarly, the associated

transformation in the vector basis reads

$$\boldsymbol{e}_{\hat{a}} = e_{\hat{a}}{}^{\mu}\partial_{\mu} \iff \partial_{\mu} = e_{\mu}{}^{\hat{a}}\boldsymbol{e}_{\hat{a}}. \tag{10.17}$$

From Eq. (10.16), we can write the metric $g_{\mu\nu}$ in the new basis:

$$g_{\hat{a}\hat{b}} = g_{\mu\nu}e^{\mu}{}_{\hat{a}}e^{\nu}{}_{\hat{b}}. \tag{10.18}$$

A special choice of basis corresponds to the metric in the new frame being the Minkowski metric $g_{\hat{a}\hat{b}} = \eta_{\hat{a}\hat{b}} = \mathrm{diag}\,(-1,1,1,1)$. The corresponding basis $e^{\hat{a}}$ is called an orthonormal frame. There again, in general, the new basis cannot be a coordinate basis, otherwise the metric of the spacetime would be equivalent to the Minkowski metric by coordinate transformation and spacetime would be flat. Also, note that such orthonormal bases are not unique: they are defined up to a Lorentz transformation, which preserves the Minkowski metric by definition.

This hints toward a procedure to apply the flat spacetime Boris pusher. The components of the momentum u_i and the fields B^i, D^i are defined in the spherical coordinate basis ∂_i. They are initially transformed to the orthonormal basis $\boldsymbol{e}_{\hat{a}}$ and $e^{\hat{a}}$, with $\hat{a} = x$, y or z, according to

$$u_{\hat{a}}(t) = e_{\hat{a}}{}^{i}u_i(t), \quad B^{\hat{a}} = e^{\hat{a}}{}_{i}B^i, \quad D^{\hat{a}} = e^{\hat{a}}{}_{i}D^i. \tag{10.19}$$

Then the Boris push is applied using the orthonormal coordinates: $u_{\hat{a}}(t) \to u_{\hat{a}}(t + \Delta t/2)$. At the end of the pusher, the components of the momentum are transformed back to the coordinate basis with $u_i\,(t + \Delta t/2) = e^{\hat{a}}{}_{i}u_{\hat{a}}(t+\Delta t/2)$. The radiation reaction force, if present, is computed and added to the dynamics of the particle the same way. The orthonormal tetrads in KS coordinates and the transformation of 3-vectors and 1-forms, which are used in the code, are given in the Appendix 10.B.

10.1.5 Charge deposition

The charge deposition scheme is similar to the special relativistic version, and we use a 2D linear interpolation scheme. The proportion of a charge which is deposited on a given node is is determined by the volume defined by the position of the particle in the cell (see Fig. 10.2) (Cerutti et al., 2015, 2016). If a macroparticle with charge q and weight w is located between the nodes $r \in [r_i, r_{i+1}]$ and $\theta \in [\theta_j, \theta_{j+1}]$, then the charge deposited on each node is

$$q_{i,j} = \frac{V_{i+1,j+1}}{V}\,q\,w, \quad q_{i,j+1} = \frac{V_{i+1,j}}{V}\,q\,w, \tag{10.20}$$

$$q_{i+1,j} = \frac{V_{i,j+1}}{V}\,q\,w, \quad q_{i+1,j+1} = \frac{V_{i,j}}{V}\,q\,w. \tag{10.21}$$

where the volume of the cell V is

$$V = \int_{r_i}^{r_{i+1}}\int_{\theta_j}^{\theta_{j+1}} \sqrt{h}\,\mathrm{d}r\,\mathrm{d}\theta. \tag{10.22}$$

The volume $V_{i,j}$, for instance, reads

$$V_{i,j} = \int_{r_i}^{r}\int_{\theta_j}^{\theta} \sqrt{h}\,\mathrm{d}r\,\mathrm{d}\theta. \tag{10.23}$$

The different volumes can be evaluated by numerical integration.

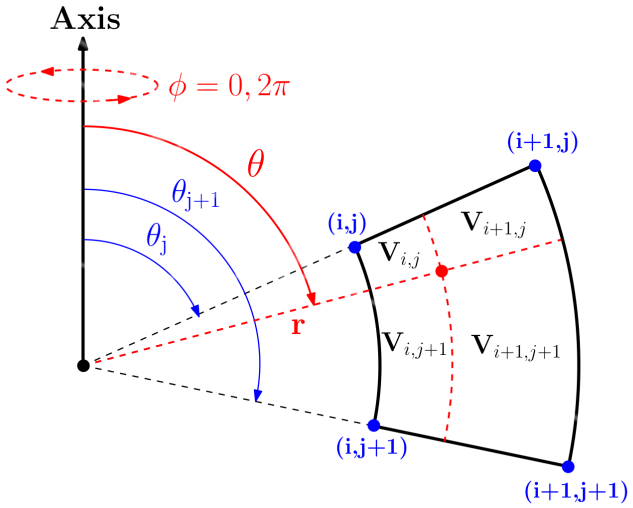


FIGURE 10.2: Volume weighting procedure in spherical coordinates. Figure taken from Cerutti et al. (2015).

10.2 Electromagnetic solver

10.2.1 Spatial discretization

The staggering of the different components of the fields is different in spherical coordinates. A spherical Yee grid for special relativistic simulations was designed independently by Cerutti et al. (2015) and Belyaev (2015a), and by Cerutti et al. (2016) in 3D. It can be generalized to encompass the 4 fields $\boldsymbol{D}$, $\boldsymbol{B}$, $\boldsymbol{E}$ and $\boldsymbol{H}$, as sketched in Fig. 10.3.

Working in the coordinate basis has its perks: the discrete forms of the curl and divergence of a vector field are easier to derive and write down in that basis than in the orthonormal spherical basis (see the intricate expressions in Cerutti et al. (2015)). Since the quantity $\sqrt{h}$ is often involved in these expressions, it must be stored by the code separately on every node of the spherical Yee mesh. For illustrative purposes, we provide here the discretized expression of the different components of $(\boldsymbol{\nabla} \times \boldsymbol{E})^i = \varepsilon^{ijk} \partial_j E_k / \sqrt{h}$, denoting Δr and $\Delta\theta$ the local grid spacing in r and θ:

$$(\boldsymbol{\nabla} \times \boldsymbol{E})^r_{i,j+1/2} = \frac{(E_\varphi)_{i,j+1} - (E_\varphi)_{i,j}}{\Delta\theta \sqrt{h}_{i,j+1/2}}, \tag{10.24}$$

$$(\boldsymbol{\nabla} \times \boldsymbol{E})^\theta_{i+1/2,j} = -\frac{(E_\varphi)_{i+1,j} - (E_\varphi)_{i,j}}{\Delta r \sqrt{h}_{i+1/2,j}}, \tag{10.25}$$

$$(\boldsymbol{\nabla} \times \boldsymbol{E})^\varphi_{i+1/2,j+1/2} = \frac{(E_\theta)_{i+1,j+1/2} - (E_\theta)_{i,j+1/2}}{\Delta r \sqrt{h}_{i+1/2,j+1/2}} - \frac{(E_r)_{i+1/2,j+1} - (E_\theta)_{i+1/2,j}}{\Delta\theta \sqrt{h}_{i+1/2,j+1/2}}. \tag{10.26}$$

Note that we reserve the possibility of using grids unevenly spaced in r and θ. It can be appropriate to use a radial grid equally spaced in $\log(r)$ to increase the size of the simulation domain at a reduced computational cost, or an orthoradial grid equally spaced in $\cos\theta$ to increase the resolution in the equatorial plane. In that case, the spacing Δr and $\Delta\theta$ does not remain constant. In the above equations and in the following, it is implicit that Δr and $\Delta\theta$ are centered on the computed node. For instance, we use $\Delta\theta_{j+1/2} = \theta_{j+1} - \theta_j$ in Eq. (10.24) but $\Delta\theta_j = \theta_{j+1/2} - \theta_{j-1/2}$ in Eq. (10.26). Besides, in KS coordinates, the determinant h of the spatial metric reads $\sqrt{h} = \Sigma \sin\theta / \alpha$. This can lead to divergences close to the axis, as

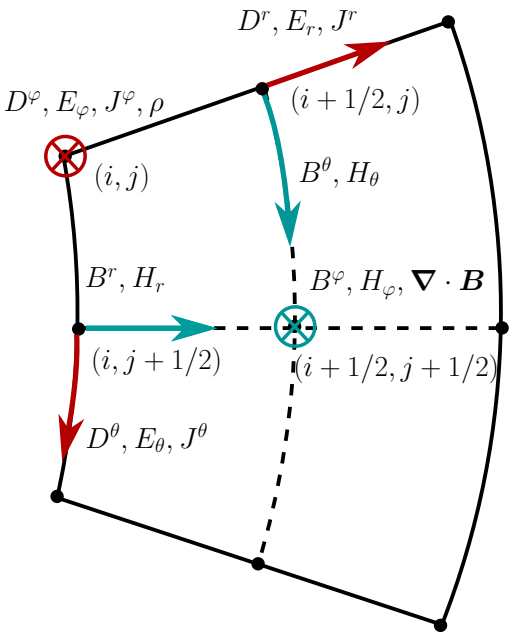


FIGURE 10.3: Yee grid in spherical coordinates. The red arrows denote the components of the electric fields, the blue ones denote the magnetic fields. The positions on the mesh are labelled by the indices i in the radial direction and j in the orthoradial direction.

$\sin\theta \to 0$. For this reason, the denominator in Eq. (10.24) for instance is rewritten in the code as $\sqrt{h}\Delta\theta = \sqrt{h'}|\Delta\cos\theta|$, with $h' = h/\sin^2\theta$.

This staggering on the spherical Yee grid automatically enforces the conservation of magnetic flux $\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0$, just like in the Cartesian case.

10.2.2 Time evolution

The code evolves Maxwell's equations (5.37b) and (5.37d) for the fields $\boldsymbol{B}$ and $\boldsymbol{D}$, supplemented with the constitutive relations (5.41a) and (5.41b) which prescribe $\boldsymbol{E}$ and $\boldsymbol{H}$. These relations come with their own challenges, since the 4 fields must be evolved while keeping a time-centered leapfrog architecture. The algorithm designed by Parfrey et al. (2019) goes as follows. Let us assume that $\boldsymbol{B}$ is known at the time steps $n-1$ and n, and $\boldsymbol{D}$ is known at the time steps $n-1/2$ and $n+1/2$. Our goal is first to derive $\boldsymbol{B}^{n+1}$.

1. Compute the means $\boldsymbol{B}^{n-1/2} \simeq (\boldsymbol{B}^{n-1} + \boldsymbol{B}^n)/2$ and $\boldsymbol{D}^n \simeq (\boldsymbol{D}^{n-1/2} + \boldsymbol{D}^{n+1/2})/2$.
2. $\boldsymbol{E}^n = \alpha\boldsymbol{D}^n + \boldsymbol{\beta}\times\boldsymbol{B}^n$.
3. $\boldsymbol{B}^{n-1/2} \to \boldsymbol{B}^{n+1/2}$ with $(\boldsymbol{\nabla}\times\boldsymbol{E})^n$.
4. $\boldsymbol{E}^{n+1/2} = \alpha\boldsymbol{D}^{n+1/2} + \boldsymbol{\beta}\times\boldsymbol{B}^{n+1/2}$.
5. $\boldsymbol{B}^n \to \boldsymbol{B}^{n+1}$ with $(\boldsymbol{\nabla}\times\boldsymbol{E})^{n+1/2}$.

Now the second step consists in pushing the particles, in order to get the electric current density needed to evolve $\boldsymbol{D}$.

6. Compute the arithmetical mean $\boldsymbol{B}^{n+1/2}$.
7. ${x^i}^n \to {x^i}^{n+1}$ and ${u_i}^n \to {u_i}^{n+1}$, using $\boldsymbol{B}^{n+1/2}$ and $\boldsymbol{D}^{n+1/2}$ in the two runs of the Boris push.

8. Compute the updated source terms $\rho\left({x^i}^{n+1}, {u_i}^{n+1}\right)$, $\boldsymbol{J}\left({x^i}^{n+1}, {u_i}^{n+1}\right)$.

Now the $\boldsymbol{D}$ field can be computed, through steps very similar to the five first steps for $\boldsymbol{B}$.

9. Compute the arithmetical mean $\boldsymbol{J}^{n+1/2}$.
10. $\boldsymbol{H}^{n+1/2} = \alpha \boldsymbol{B}^{n+1/2} - \boldsymbol{\beta} \times \boldsymbol{D}^{n+1/2}$.
11. $\boldsymbol{D}^n \to \boldsymbol{D}^{n+1}$ with $(\boldsymbol{\nabla} \times \boldsymbol{H})^{n+1/2}$ and $\boldsymbol{J}^{n+1/2}$.
12. $\boldsymbol{H}^{n+1} = \alpha \boldsymbol{B}^{n+1} - \boldsymbol{\beta} \times \boldsymbol{D}^{n+1}$.
13. $\boldsymbol{D}^{n+1/2} \to \boldsymbol{D}^{n+3/2}$ with $(\boldsymbol{\nabla} \times \boldsymbol{H})^{n+1}$ and $\boldsymbol{J}^{n+1}$.

Preserving the leapfrog sheme has come at a cost: twice as many leapfrog pushes are needed in order to evolve the fields, compared to the special relativistic version. Indeed, in this case, the fields must be computed for all integer and half-integer time steps. In addition, several computations of the auxiliary fields must be performed. This algorithm is more expensive computationally. It reduces to the standard leapfrog scheme in the limit $\alpha = 1$, $\boldsymbol{\beta} = \boldsymbol{0}$, with a time step $\Delta t/2$ instead of Δt. Besides, the fields $\boldsymbol{B}$ and $\boldsymbol{D}$ must be stored for two successive time steps, occupying more memory space.

Several other hidden steps are involved. To illustrate this, let us look at the evolution equation for B^r, which reads

$$\frac{(B^r)^{n+1}_{i,j+1/2} - (B^r)^n_{i,j+1/2}}{c\Delta t} = \frac{(E_\varphi)^n_{i,j+1} - (E_\varphi)^n_{i,j}}{|\Delta \cos\theta| \sqrt{h'}_{i,j+1/2}}. \tag{10.27}$$

In this push, the value of E_φ can be obtained by computing

$$E_\varphi = \alpha D_\varphi + \sqrt{h}\beta^r B^\theta. \tag{10.28}$$

The first subtlety lies in the fact that $\boldsymbol{B}$ and $\boldsymbol{E}$ are not located at the same nodes on the Yee grid. To preserve second-order accuracy, the magnetic field is defined by a centered spatial average. In the previous example, B^θ (defined at $(i+1/2, j)$) is computed at (i, j) using

$${B^\theta}_{i,j} = \frac{\sqrt{h}_{i+1/2,j} {B^\theta}_{i+1/2,j} - \sqrt{h}_{i-1/2,j} {B^\theta}_{i-1/2,j}}{\sqrt{h}_{i+1/2,j} + \sqrt{h}_{i-1/2,j}}. \tag{10.29}$$

The weights $\sqrt{h}$ are necessary in order to keep second-order accuracy.

Besides, the covariant components of $\boldsymbol{E}$ are needed to evolve $\boldsymbol{B}$, yet only the contravariant components of $\boldsymbol{D}$ are stored and evolved. In the previous example, the covariant component D_φ must first be computed from D^i with $D_i = h_{ij} D^j$:

$$D_\varphi = h_{\varphi\varphi} D^\varphi + h_{r\varphi} D^r. \tag{10.30}$$

Recall that in the KS coordinates, h_{ij} is not diagonal and β is purely radial. The same issue can be raised in Eq. (10.30): D^r is defined on the node $(i+1/2, j)$ instead of (i, j), so it must also be averaged before the computation of $D_{\varphi i,j}$.

10.2.3 Divergence cleaning

Like in the special-relativistic version, Maxwell-Gauss equation $\boldsymbol{\nabla} \cdot \boldsymbol{D} = 4\pi\rho$ is not enforced by the field pusher, so divergence cleaning must be periodically performed. Poisson's equation $\Delta\phi = \boldsymbol{\nabla} \cdot \boldsymbol{D} - 4\pi\rho$ must be solved. Using the general definition of the gradient and the divergence, the Laplacian of ϕ (defined at the nodes (i, j)) reads

$$\Delta\phi = \frac{1}{\sqrt{h}}\partial_k\left(\sqrt{h}h^{kl}\partial_l\phi\right). \tag{10.31}$$

The source term on the right of Poisson's equation is discretized as

$$S_{i,j} = \frac{\left(D^r\sqrt{h}\right)_{i+1/2,j} - \left(D^r\sqrt{h}\right)_{i-1/2,j}}{\Delta r\sqrt{h}_{i,j}} + \frac{\left(D^\theta\sqrt{h}\right)_{i,j+1/2} - \left(D^\theta\sqrt{h}\right)_{i,j-1/2}}{\Delta\theta\sqrt{h}_{i,j}} - 4\pi\rho_{i,j}. \tag{10.32}$$

Noting $(\nabla\phi)_i = \partial_i\phi$ the covariant components of the gradient of ϕ, we can discretize the Laplacian as

$$(\Delta\phi)_{i,j} = \frac{1}{\sqrt{h}_{i,j}}\frac{\left(\sqrt{h}\,(\nabla\phi)^r\right)_{i+1/2,j} - \left(\sqrt{h}\,(\nabla\phi)^r\right)_{i-1/2,j}}{\Delta r} \tag{10.33}$$

$$+ \frac{1}{\sqrt{h}_{i,j}}\frac{\left(\sqrt{h}\,(\nabla\phi)^\theta\right)_{i,j+1/2} - \left(\sqrt{h}\,(\nabla\phi)^r\right)_{i,j-1/2}}{\Delta\theta}. \tag{10.34}$$

Discretizing the gradient, with $(\nabla\phi)^r_{i+1/2,j} = h^{rr}_{i+1/2,j}\left(\phi_{i+1,j} - \phi_{i,j}\right)/\Delta r$ for instance, leads to

$$(\Delta\phi)_{i,j} = \frac{1}{\sqrt{h}_{i,j}}\left(C_1\left(\phi_{i+1,j} - \phi_{i,j}\right) + C_2\left(\phi_{i-1,j} - \phi_{i,j}\right) + C_3\left(\phi_{i,j+1} - \phi_{i,j}\right) + C_4\left(\phi_{i,j-1} - \phi_{i,j}\right)\right), \tag{10.35}$$

where the weight coefficients area

$$C_1 = \frac{\sqrt{h}_{i+1/2,j}h^{rr}_{i+1/2,j}}{\Delta r^2}, \quad C_2 = \frac{\sqrt{h}_{i-1/2,j}h^{rr}_{i-1/2,j}}{\Delta r^2}, \tag{10.36}$$

$$C_3 = \frac{\sqrt{h}_{i,j+1/2}h^{\theta\theta}_{i,j+1/2}}{\Delta\theta^2}, \quad C_4 = \frac{\sqrt{h}_{i,j-1/2}h^{\theta\theta}_{i,j-1/2}}{\Delta\theta^2}. \tag{10.37}$$

Eventually, we can express implicitly $\phi_{i,j}$ as

$$\phi_{i,j} = \frac{\sqrt{h}_{i,j}S_{i,j}}{C_1 + C_2 + C_3 + C_4} + \frac{C_1\phi_{i+1,j} + C_2\phi_{i-1,j} + C_3\phi_{i,j+1} + C_4\phi_{i,j-1}}{C_1 + C_2 + C_3 + C_4}. \tag{10.38}$$

Although the discretization is slightly different, the iterative Jacobi method can again be used to find $\phi_{i,j}$, so the corrected electric field becomes $\boldsymbol{D} = \boldsymbol{D}_0 - \boldsymbol{\nabla}\phi$.

10.2.4 CFL condition

In order to derive the CFL condition in this geometry, the same method than in the Cartesian geometry can be applied. Let us consider a single Fourier component of the fields of frequency ω and wave vector $\boldsymbol{k}$, which reads

$$\boldsymbol{X}^n_{l,m} = \boldsymbol{X}_0\exp\left(i\left(\omega n\Delta t + lk_r\Delta r + mk_\theta\Delta\theta\right)\right). \tag{10.39}$$

The covariant components of the wave vector are involved in the scalar product $\boldsymbol{k}\cdot\boldsymbol{x} = k_i x^i$. Our goal is to derive the dispersion relation for a light wave in vacuum. Inserting Eq. (10.39) into Eq. (5.37d), we find the discretized equation

$$\Omega\boldsymbol{B} = -\boldsymbol{K}\times\boldsymbol{E}, \tag{10.40}$$

where

$$\Omega = \frac{\sin\omega\Delta t/2}{c\Delta t/2}, \quad K_i = \frac{\sin k_i\Delta x^i/2}{\Delta x^i/2} = k_i\frac{\sin k_i\Delta x^i/2}{k_i\Delta x^i/2}. \tag{10.41}$$

The "discretized" K_i components transform as a 1-form. In this equation, the phase factors have been absorbed in the complex amplitude of the fields. Similarly, Maxwell-Ampère's equation is transformed into $\Omega\boldsymbol{D} = \boldsymbol{K}\times\boldsymbol{H}$.

From the previous discretized equations for $\boldsymbol{B}$ and $\boldsymbol{D}$, it follows that $\boldsymbol{K}\cdot\boldsymbol{B} = \boldsymbol{K}\cdot\boldsymbol{D} = 0$. Now, combining Eq. (10.40) with the constitutive relation defining $\boldsymbol{E}$ (see Eq. (5.41a)) we obtain $\Omega\boldsymbol{B} = -\alpha\boldsymbol{K}\times\boldsymbol{D} - \boldsymbol{K}\times(\boldsymbol{\beta}\times\boldsymbol{B})$. Expanding the double cross product with $\boldsymbol{K}\cdot\boldsymbol{B} = 0$ leads to

$$(\Omega - \boldsymbol{K}\cdot\boldsymbol{\beta})\,\boldsymbol{B} = -\alpha\boldsymbol{K}\times\boldsymbol{D}. \tag{10.42}$$

Similarly, the other equation $\Omega\boldsymbol{D} = \boldsymbol{K}\times\boldsymbol{H}$ can be combined with the constitutive relation (5.41b) into $(\Omega - \boldsymbol{K}\cdot\boldsymbol{\beta})\,\boldsymbol{D} = \alpha\boldsymbol{K}\times\boldsymbol{B}$. All in all, we get

$$(\Omega - \boldsymbol{K}\cdot\boldsymbol{\beta})^2\,\boldsymbol{B} = -\alpha^2\boldsymbol{K}\times(\boldsymbol{K}\times\boldsymbol{B}) = \alpha^2\left(K_iK^i\right)\boldsymbol{B}. \tag{10.43}$$

Hence, the dispersion relation can be put into the following form:

$$\left(\frac{\sin\omega\Delta t/2}{c\alpha\Delta t/2} - \frac{\beta^i}{\alpha}\frac{\sin k_i\Delta x^i/2}{\Delta x^i/2}\right)^2 = h^{rr}\left(\frac{\sin k_r\Delta r/2}{\Delta r/2}\right)^2 + h^{\theta\theta}\left(\frac{\sin k_\theta\Delta\theta/2}{\Delta\theta/2}\right)^2. \tag{10.44}$$

The CFL condition finally reads, with $\boldsymbol{\beta} = \beta^r\partial_r$ in KS coordinates:

$$\frac{1}{c\alpha\Delta t} < \frac{\beta^r}{\alpha\Delta r} + \sqrt{\frac{h^{rr}}{\Delta r^2} + \frac{h^{\theta\theta}}{\Delta\theta^2}}. \tag{10.45}$$

In flat spacetime and with $\boldsymbol{\beta} = 0$, this reduces to the more standard CFL condition in spherical coordinates (Cerutti et al., 2015).

10.A Appendix: Derivation of the $3+1$ equations of motion

10.A.1 Neutral particle

The Lagrangian of a particle (massive or massless) in general relativity reads (Blau, 2018)

$$\mathcal{L} = \frac{1}{2}g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}, \tag{10.46}$$

where λ parametrizes the wordline. For a massive particle, λ can be taken to be the proper time along the wordline. The Euler-Lagrange equations with this choice of Lagrangian indeed yield the famous geodesic equation

$$\frac{\mathrm{d}^2x^\mu}{\mathrm{d}\lambda^2} + \Gamma^\mu{}_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda}\frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda} = \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda}\nabla_\nu\frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} = 0. \tag{10.47}$$

Because by definition, the tangent vector $\dot{x}^\mu = \mathrm{d}x^\mu/\,\mathrm{d}\lambda$ is parallel-transported along a geodesic, the Lagrangian (10.46) possesses the pleasant property of being conserved along a geodesic. This preserves the nature (timelike, spacelike or null) of the wordline. As a consequence, one can choose a suitable parametrization so as to impose

$$g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu = \text{cst} \tag{10.48}$$

at all times, the constant being zero for a massless particle. The associated action is

$$S\left[x^\mu\left(\lambda\right)\right] = \int \frac{1}{2} g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu \,\mathrm{d}\lambda. \tag{10.49}$$

The canonical momentum conjugated to x^μ are $p_\mu = \partial\mathcal{L}/\dot{x}^\mu = g_{\mu\nu}\dot{x}^\nu$, and the Hamiltonian is defined by the Legendre transformation

$$\mathcal{H}\left(x^\mu, p_\nu\right) = p_\mu \frac{\mathrm{d}x^\mu}{\mathrm{d}\tau} - \mathcal{L} = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu. \tag{10.50}$$

The Hamiltonian is conserved and actually does not depend on λ. It must be viewed as a function of the phase space coordinates of the spacetime (x^μ, p_ν).

These Lagrangian, action and Hamiltonian are manifestly covariant. However, they are quite inconvenient in the sense that every particle has its own affine parameter λ. We would like to have a single time $t = x^0/c$, our global universal coordinate time in the $3+1$ formalism, to parametrize all geodesics. We need to define a Hamiltonian which is a function of $\left(x^i, p_j\right)$, the coordinates of the phase space restricted to a hypersurface of the foliation. Actually, the action (10.49) can be rewritten as a functional of the phase space trajectory in the hypersurface:

$$S\left[x^j(t), p_i(t)\right] = \frac{1}{2}\int p_\mu \,\mathrm{d}x^\mu = \frac{1}{2}\int \left(cp_0 + p_i \frac{\mathrm{d}x^i}{\mathrm{d}t}\right) \mathrm{d}t. \tag{10.51}$$

The new Lagrangian is $L = cp_0 + \mathrm{d}x^i/\,\mathrm{d}t$. Another Legendre transformation yields the Hamiltonian of the particle: $\mathcal{H}\left(x^i, p_i\right) = -cp_0$. This quantity is actually the conserved energy, arising from time translational invariance (see Eq. (4.21)).

Let us now express this Hamiltonian as a function of x^i and p_i. According to Eq. (10.48), the affine parameter is defined so that the following relation holds along the geodesic:

$$g^{\mu\nu} p_\mu p_\nu = -m^2c^2\varepsilon, \tag{10.52}$$

where $\varepsilon = 1$ if the particle has mass m, and $\varepsilon = 0$ if the particle is massless. Expanding this equation, we find in the $3+1$ formalism:

$${p_0}^2 - 2\left(\beta^i p_i\right) p_0 + \left(\beta^i p_i\right)^2 - \alpha^2 h^{jk} p_j p_k - \alpha^2 m^2 c^2 \varepsilon = 0. \tag{10.53}$$

This is a quadratic equation in p_0, which allows us to write p_0 as a function of p_i. In the case of a massive particle ($\varepsilon = 1$), the reduced discriminant is $\alpha^2 m^2 c^2 \Gamma^2$, where we have defined the normalized momentum $u_i = p_i/mc$ and the Lorentz factor of the particle $\Gamma = \sqrt{1 + h^{jk}u_j u_k}$. The sign of the solution is determined by the flat-spacetime limit ($\beta^i \to 0$, $\alpha \to 1$). All in all, the $3+1$ Hamiltonian of a massive particle reads

$$\mathcal{H}\left(x^j, p_i\right) = -cp_0 = \alpha mc^2 \sqrt{1 + h^{jk}\frac{p_j p_k}{m^2c^2}} - c\beta^i p_i = mc^2\left(\alpha\Gamma - \beta^i u_i\right). \tag{10.54}$$

Now we can write down Hamilton's equations for the evolution of the particle coordinates x^i and normalized momentum u_i:

$$\frac{1}{c}\frac{\mathrm{d}x^i}{\mathrm{d}t} = \frac{1}{c}\frac{\partial \mathcal{H}}{\partial p_i} = \frac{\alpha}{\Gamma}h^{ij}u_j - \beta^i, \tag{10.55}$$

$$\frac{1}{c}\frac{\mathrm{d}u_i}{\mathrm{d}t} = -\frac{1}{mc^2}\frac{\partial \mathcal{H}}{\partial x^i} = -\Gamma\partial_i\alpha + u_j\partial_i\beta^j - \frac{\alpha}{2\Gamma}u_j u_k \partial_i h^{jk}. \tag{10.56}$$

This procedure is valid only if extremizing the action in Eq. (10.49) with respect to all possible wordlines $x^\mu(\lambda)$, for an affine parameter λ, is equivalent to extremizing the action in Eq. (10.51) with respect to all trajectories in phase space $\left(x^i(t), p_i(t)\right)$. The proof is rather involved, and can be found in Bertschinger (1999).

10.A.2 Charged particle

We include the effect of an electromagnetic field with 4-potential A_μ by the means of a minimal coupling procedure. Minimal coupling prescribes the following conjugate momentum for a particle of charge q:

$$\Pi_\mu = p_\mu + \frac{q}{c}A_\mu = mc\left(u_\mu + \frac{q}{mc^2}A_\mu\right). \tag{10.57}$$

The normalized conjugate momentum is defined as $\pi_\mu = \Pi_\mu/mc$. By minimal coupling, the Hamiltonian of the charged particle is

$$\mathcal{H}\left(x^i, \Pi_j\right) = -c\Pi_0 = -cp_0 - qA_0, \tag{10.58}$$

where $-cp_0$ is the Hamiltonian that we have already derived (Eq. (10.54)). In terms of $\Pi_i = mc\pi_i$, the $3+1$ Hamiltonian reads

$$\mathcal{H}(x^i, \Pi_j) = mc^2\left(\alpha\sqrt{1 + h^{ij}\left(\pi_i - \frac{q}{mc^2}A_i\right)\left(\pi_j - \frac{q}{mc^2}A_j\right)} - \beta^i\left(\pi_i - \frac{q}{mc^2}A_i\right) - \frac{q}{mc^2}A_0\right). \tag{10.59}$$

This Hamiltonian reduces to (10.54) for $A_\mu = 0$. The equations of motion are again derived from Hamilton's equations, reinserting $u_i = \pi_i - (q/mc^2)A_i$ after derivating:

$$\frac{1}{c}\frac{\mathrm{d}x^i}{\mathrm{d}t} = \frac{\alpha}{\Gamma}h^{ij}u_j - \beta^i, \tag{10.60}$$

$$\frac{1}{c}\frac{\mathrm{d}\pi_i}{\mathrm{d}t} = -\Gamma\partial_i\alpha + u_j\partial_i\beta^j - \frac{\alpha}{2\Gamma}u_j u_k\partial_i h^{jk} + \frac{q}{mc^2}\left(\partial_i A_0 + \frac{\alpha}{\Gamma}h^{jk}u_j\partial_i A_k - \beta^k\partial_i A_k\right). \tag{10.61}$$

The first equation is the same as before. The time derivatives of u_i and π_i are related by

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} = \frac{\mathrm{d}\pi_i}{\mathrm{d}t} - \frac{q}{mc^2}\partial_t A_i - \frac{q}{mc^2}\frac{\mathrm{d}x^j}{\mathrm{d}t}\partial_j A_i = \frac{\mathrm{d}\pi_i}{\mathrm{d}t} - \frac{q}{mc^2}\partial_t A_i - \frac{q}{mc^2}\frac{\alpha}{\Gamma}h^{jk}u_k\partial_j A_i + \frac{q}{mc^2}\beta^k\partial_k A_i. \tag{10.62}$$

The time derivative of u_i therefore reads

$$\frac{1}{c}\frac{\mathrm{d}u_i}{\mathrm{d}t} = -\Gamma\partial_i\alpha + u_j\partial_i\beta^j - \frac{\alpha}{2\Gamma}u_ju_k\partial_ih^{jk} + \frac{q}{mc^2}\left(\partial_iA_0 - \partial_0A_i + \frac{\alpha}{\Gamma}h^{jk}u_j\left(\partial_iA_k - \partial_kA_i\right) - \beta^k\left(\partial_iA_k - \partial_kA_i\right)\right). \tag{10.63}$$

We recognize the components of the Maxwell tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. Besides, the last term can be removed by rewriting the term $h^{jk}u_k$. The contravariant time component of u_μ reads $u^0 = -u_0/\alpha^2 + \beta^iu_i/\alpha^2 = \Gamma/\alpha$, according to Eq. (10.54). Besides, the spatial contravariant components are $u^i = g^{ij}u_j = h^{ij}u_j - \beta^i\beta^ju_j/\alpha^2 + \beta^iu_0/\alpha^2$. Inserting the expression of u_0 yields $u^i = h^{ij}u_j - \beta^i\Gamma/\alpha$ (recall that u^μ is not a 3-vector on a hypersurface of the foliation). All in all, we have

$$\frac{1}{c}\frac{\mathrm{d}u_i}{\mathrm{d}t} = -\Gamma\partial_i\alpha + u_j\partial_i\beta^j - \frac{\alpha}{2\Gamma}u_ju_k\partial_ih^{jk} + \frac{1}{mc^2}\mathcal{F}_i, \quad \mathcal{F}_i = q\frac{\alpha}{\Gamma}F_{i\mu}u^\mu, \tag{10.64}$$

where $\boldsymbol{\mathcal{F}}$ is the Lorentz force exerted by the fields on the particle. This expression of $\boldsymbol{\mathcal{F}}$ is very reminiscent of the special-relativistic Lorentz force. All that remains is to rearrange it in terms of the electromagnetic fields $\boldsymbol{D}$ and $\boldsymbol{B}$. Using $B^i = e^{ijk}F_{jk}/2 \Rightarrow F_{ij} = e_{ijk}B^k$ and $u^0 = \Gamma/\alpha$, we have

$$\begin{aligned} F_{i\mu}u^\mu &= F_{i0}u^0 + F_{ij}u^j && (10.65)\\ &= E_iu^0 + e_{ijk}u^jB^k && (10.66)\\ &= \left(\alpha D_i + e_{ijk}\beta^jB^k\right)u^0 + e_{ijk}\left(h^{jl}u_l - u^0\beta^j\right)B^k && (10.67)\\ &= \Gamma D_i + e_{ijk}h^{jl}u_lB^k. && (10.68)\end{aligned}$$

This finally yields the expression of the Lorentz force as a function of the fields measured by the FIDO:

$$\mathcal{F}_i = q\alpha\left(h_{ij}D^j + e_{ijk}\frac{h^{jl}u_l}{\Gamma}B^k\right). \tag{10.69}$$

10.A.3 Massless particle

The derivation is completely similar for the gedoesic of a photon (massless particle), removing the Lorentz force and enforcing

$$g^{\mu\nu}p_\mu p_\nu = 0. \tag{10.70}$$

In this case, the $3+1$ Hamiltonian reads

$$\mathcal{H}\left(x^j, p_i\right) = -cp_0 = c\alpha\sqrt{h^{jk}p_jp_k} - c\beta^ip_i. \tag{10.71}$$

We define $\mathcal{E} = c\sqrt{h^{jk}p_jp_k}$ as the FIDO-measured energy of the photon. The equations of motion of the photon are

$$\begin{aligned} \frac{1}{c}\frac{\mathrm{d}x^i}{\mathrm{d}t} &= c\frac{\alpha}{\mathcal{E}}h^{ij}p_j - \beta^i, && (10.72)\\ \frac{1}{c}\frac{\mathrm{d}p_i}{\mathrm{d}t} &= -\frac{1}{c}\mathcal{E}\partial_i\alpha + p_j\partial_i\beta^j - \frac{c\alpha}{2\mathcal{E}}p_jp_k\partial_ih^{jk}. && (10.73)\end{aligned}$$

10.B Appendix: Orthonormal tetrads in KS coordinates

The easiest way to derive the expression of the orthonormal tetrads is to rewrite the metric $g_{\mu\nu}$ in a way that resembles the Minkowski metric, then to perform the appropriate transformation of 1-forms. Namely, according to Eq. (5.17), the line element in KS coordinates can be rewritten as

$$\begin{aligned}
\mathrm{d}s^2 &= -\alpha^2 c^2\,\mathrm{d}t^2 + h_{ij}\left(\mathrm{d}x^i + \beta^i c\,\mathrm{d}t\right)\left(\mathrm{d}x^j + \beta^j c\,\mathrm{d}t\right) && (10.74)\\
&= -\alpha^2 c^2\,\mathrm{d}t^2 + h_{rr}\left(\mathrm{d}r + \beta^r c\,\mathrm{d}t\right)^2 + h_{\theta\theta}\,\mathrm{d}\theta^2 + h_{\varphi\varphi}\,\mathrm{d}\varphi^2 + 2h_{r\varphi}\left(\mathrm{d}r + \beta^r c\,\mathrm{d}t\right)\mathrm{d}\varphi && (10.75)\\
&= -\alpha^2 c^2\,\mathrm{d}t^2 + h_{rr}\left(\mathrm{d}r + \beta^r c\,\mathrm{d}t\right)^2 + h_{\theta\theta}\,\mathrm{d}\theta^2 \\
&\quad + h_{\varphi\varphi}\left(\mathrm{d}\varphi + \frac{h_{r\varphi}}{h_{\varphi\varphi}}\left(\mathrm{d}r + \beta^r c\,\mathrm{d}t\right)\right)^2 - \frac{{h_{r\varphi}}^2}{h_{\varphi\varphi}}\left(\mathrm{d}r + \beta^r c\,\mathrm{d}t\right)^2 && (10.76)\\
&= -\alpha^2 c^2\,\mathrm{d}t^2 + h_{\theta\theta}\,\mathrm{d}\theta^2 + \frac{1}{\mathcal{A}^2}\left(\mathrm{d}r + \beta^r c\,\mathrm{d}t\right)^2 + h_{\varphi\varphi}\left(\mathrm{d}\varphi + \frac{h_{r\varphi}}{h_{\varphi\varphi}}\left(\mathrm{d}r + \beta^r c\,\mathrm{d}t\right)\right)^2. && (10.77)
\end{aligned}$$

We have introduced the convenient notation $\mathcal{A} = \sqrt{h_{\varphi\varphi}/\left(h_{rr}h_{\varphi\varphi} - {h_{r\varphi}}^2\right)} = \sqrt{h^{rr}}$. The last equality comes from $h = h_{\theta\theta}\left(h_{rr}h_{\varphi\varphi} - {h_{r\varphi}}^2\right)$ and Cramer's rule $h^{rr} = h_{\theta\theta}h_{\varphi\varphi}/h$. Consequently, by defining $e^{\hat{t}} = \alpha\,\mathrm{d}t$, $e^{\hat{x}} = \left(\mathrm{d}r + \beta^r c\,\mathrm{d}t\right)/\mathcal{A}$, $e^{\hat{y}} = \sqrt{h_{\theta\theta}}\,\mathrm{d}\theta$ and $e^{\hat{z}} = \sqrt{h_{\varphi\varphi}}\,\mathrm{d}\varphi + \left(h_{r\varphi}/\sqrt{h_{\varphi\varphi}}\right)\left(\mathrm{d}r + \beta^r c\,\mathrm{d}t\right)$, one indeed finds the Minkowski metric

$$\mathrm{d}s^2 = -c^2 {e^{\hat{t}}}^2 + {e^{\hat{x}}}^2 + {e^{\hat{y}}}^2 + {e^{\hat{z}}}^2. \tag{10.78}$$

As a result, an orthonormal basis is given in KS coordinates by:

$$\begin{aligned}
\boldsymbol{e}_{\hat{t}} &= \frac{1}{\alpha}\left(\partial_t - c\beta^r\partial_r\right), && \partial_t = \alpha\boldsymbol{e}_{\hat{t}} + \frac{c\beta^r}{\mathcal{A}}\boldsymbol{e}_{\hat{x}} + c\beta^r\frac{h_{r\varphi}}{\sqrt{h_{\varphi\varphi}}}\boldsymbol{e}_{\hat{z}}, && (10.79)\\
\boldsymbol{e}_{\hat{x}} &= \mathcal{A}\left(\partial_r - \frac{h_{r\varphi}}{h_{\varphi\varphi}}\partial_\varphi\right), && \partial_r = \frac{1}{\mathcal{A}}\boldsymbol{e}_{\hat{x}} + \frac{h_{r\varphi}}{\sqrt{h_{\varphi\varphi}}}\boldsymbol{e}_{\hat{z}}, && (10.80)\\
\boldsymbol{e}_{\hat{y}} &= \frac{1}{\sqrt{h_{\theta\theta}}}\partial_\theta, && \partial_\theta = \sqrt{h_{\theta\theta}}\boldsymbol{e}_{\hat{y}}, && (10.81)\\
\boldsymbol{e}_{\hat{z}} &= \frac{1}{\sqrt{h_{\varphi\varphi}}}\partial_\varphi, && \partial_\varphi = \sqrt{h_{\varphi\varphi}}\boldsymbol{e}_{\hat{z}}. && (10.82)
\end{aligned}$$

This yields the following tetrad and inverse tetrad:

$$\left({e^\mu}_{\hat{a}}\right) = \begin{pmatrix} {e^t}_{\hat{t}} & {e^t}_{\hat{x}} & {e^t}_{\hat{y}} & {e^t}_{\hat{z}} \\ {e^r}_{\hat{t}} & {e^r}_{\hat{x}} & {e^r}_{\hat{y}} & {e^r}_{\hat{z}} \\ {e^\theta}_{\hat{t}} & {e^\theta}_{\hat{x}} & {e^\theta}_{\hat{y}} & {e^\theta}_{\hat{z}} \\ {e^\varphi}_{\hat{t}} & {e^\varphi}_{\hat{x}} & {e^\varphi}_{\hat{y}} & {e^\varphi}_{\hat{z}} \end{pmatrix} = \begin{pmatrix} 1/\alpha & 0 & 0 & 0 \\ -\beta^r/\alpha & \mathcal{A} & 0 & 0 \\ 0 & 0 & 1/\sqrt{h_{\theta\theta}} & 0 \\ 0 & -\mathcal{A}h_{r\varphi}/h_{\varphi\varphi} & 0 & 1/\sqrt{h_{\varphi\varphi}} \end{pmatrix}; \tag{10.83}$$

$$\left({e_\mu}^{\hat{a}}\right) = \begin{pmatrix} {e_t}^{\hat{t}} & {e_t}^{\hat{x}} & {e_t}^{\hat{y}} & {e_t}^{\hat{z}} \\ {e_r}^{\hat{t}} & {e_r}^{\hat{x}} & {e_r}^{\hat{y}} & {e_r}^{\hat{z}} \\ {e_\theta}^{\hat{t}} & {e_\theta}^{\hat{x}} & {e_\theta}^{\hat{y}} & {e_\theta}^{\hat{z}} \\ {e_\varphi}^{\hat{t}} & {e_\varphi}^{\hat{x}} & {e_\varphi}^{\hat{y}} & {e_\varphi}^{\hat{z}} \end{pmatrix} = \begin{pmatrix} \alpha & \beta^r/\mathcal{A} & 0 & \beta^r h_{r\varphi}/\sqrt{h_{\varphi\varphi}} \\ 0 & 1/\mathcal{A} & 0 & h_{r\varphi}/\sqrt{h_{\varphi\varphi}} \\ 0 & 0 & \sqrt{h_{\theta\theta}} & 0 \\ 0 & 0 & 0 & \sqrt{h_{\varphi\varphi}} \end{pmatrix}. \tag{10.84}$$

These tetrads can be found in Komissarov (2004a) and Takahashi (2007). We simply work in the 3-basis $(e_{\hat{x}}, e_{\hat{y}}, e_{\hat{z}})$, since we only transform 3-vectors. In practice, transforming u_i and

B^i into an orthonormal frame can be done as

$$B^{\hat{x}} = e^{\hat{x}}{}_r B^r = \frac{1}{\mathcal{A}} B^r, \qquad u_{\hat{x}} = e_{\hat{x}}{}^r u_r + e_{\hat{x}}{}^\varphi u_\varphi = \mathcal{A} u_r - \mathcal{A}\frac{h_{r\varphi}}{h_{\varphi\varphi}} u_\varphi, \quad (10.85)$$

$$B^{\hat{y}} = e^{\hat{y}}{}_r B^\theta = \sqrt{h_{\theta\theta}} B^\theta, \qquad u_{\hat{y}} = e_{\hat{y}}{}^\theta u_\theta = \frac{1}{\sqrt{h_{\theta\theta}}} u_\theta, \quad (10.86)$$

$$B^{\hat{z}} = e^{\hat{z}}{}_r B^r + e^{\hat{z}}{}_\varphi B^\varphi = \frac{h_{r\varphi} B^r}{\sqrt{h_{\varphi\varphi}}} + \sqrt{h_{\varphi\varphi}} B^\varphi, \; u_{\hat{z}} = e_{\hat{z}}{}^\varphi u_\varphi = \frac{1}{\sqrt{h_{\varphi\varphi}}} u_\varphi. \quad (10.87)$$

Chapter 11

Radiative transfer module

In the first part of my thesis, I have included two of the radiative processes described in Sec. 8 into `GRZeltron`: inverse Compton scattering and photon-photon pair production. In this section, I present these radiative transfer modules, which are in some sense a generalization of the algorithm presented in Levinson & Cerutti (2018), although the specifics differ. This section is partly adapted from the supplemental material in Crinquand et al. (2020).

11.1 Principles

We assume that the magnetosphere is embedded in a soft radiation field, which is emitted by the accretion flow. High-energy particles can upscatter soft photons to very high energies, promoting them to the gamma-ray band. Besides, gamma-ray photons can annihilate with these soft photons and produce pairs. In this case, pair production occurs in a regime where the two photons have very different energies $\epsilon_1 \gg \epsilon_0$, so that the simulation is in the domain of validity of the results outlined in Sec. 8.2. Gamma-ray photons are included in the code as a neutral third species which follows null geodesics (see the Appendix 10.A.3). We assume that the plasma is optically thin, so that the only source of opacity for the photons is photon-photon annihilation.

Instead of simulating the individual photons composing the soft radiation field, we consider that electrons, positrons and high-energy photons are propagating in a continuous, opaque medium. Besides, we take advantage of the fact that FIDOs are locally inertial observers, so that the laws of special relativity can be applied, provided we only use FIDO-measured physical quantities. For simplicity, we assume that the radiation field is time-independent, uniform, isotropic and monoenergetic, with energy ϵ_0 and density n_0 in the frame of the FIDO. This is an approximation, as we ignore the redshift of photons as measured by the FIDO. Nevertheless, ignoring this physical ingredient preserves the simplicity and efficiency of the algorithm.

We do not include any feedback of the simulation on this radiation field. The upscattered photons and created leptons are assumed to propagate along the same direction as their high-energy parents, reflecting strong relativistic beaming. The fiducial optical depth of both processes is $\tau_0 = n_0 \sigma_{\mathrm{T}} r_{\mathrm{g}}$. We also neglect pairs that would be produced by the annihilation of the soft background radiation field on itself, i.e. due to MeV emission from the radiatively-inefficient flow. The density of pairs created through this process is usually expected to be much smaller than the Goldreich-Julian density, and therefore too low to screen the gap, for the very low accretion rate found for M87* (Levinson & Rieger, 2011; Mościbrodzka et al., 2011).

Let us consider a particle (lepton or photon) which propagates into this opaque medium. We have defined in Sec. 8 the optical depth τ traversed by the particle, such that the probability that the particle interacts with the medium is $1 - \exp(-\tau)$. At every time step, we compute the infinitesimal optical depth

$$\delta\tau = \kappa \delta l, \tag{11.1}$$

where κ is the opacity for inverse Compton scattering (Eq. (8.22)) if the particle is a lepton, or the opacity of photon-photon annihilation (Eq. (8.36)) if the particle is a photon. $\delta l = \sqrt{h_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j}$ is the proper length traveled by the particle whose spatial coordinates have changed by an amount $\mathrm{d}x^i$. Then, a number p is randomly drawn with uniform probability between 0 and 1. A scattering event occurs if the condition $p < 1-\exp(-\delta\tau)$ is satisfied. For a lepton, a new high-energy photon is created along the same direction of motion, and the algorithm described in Sec. 11.2 is applied to determine its energy. For a photon, a new $e^{\pm}$ pair is created along the same direction of motion, and the algorithm described in Sec. 11.3 is applied to determine its energy.

At the beginning of the simulation, the opacities $\kappa_{\rm IC}(\gamma_0)$ and $\kappa_{\gamma\gamma}(\epsilon_1)$ are computed and stored. They are used to compute, for each particle and at every time step, the optical depth traversed by the particle as a function of the Lorentz factor (for a lepton) or the energy (for a photon).

11.2 Inverse Compton scattering

11.2.1 Kinematics

We consider, in the FIDO frame, an ultra-relativistic lepton of energy $\gamma_0 m_{\rm e}c^2$ ($\gamma_0 \gg 1$) interacting with a soft photon of energy ϵ_0. The photon makes angles (θ_0, φ_0) with the lepton velocity. We employ the same notations as in Sec. 8.1. In particular, quantities defined in the lepton rest frame will be primed. After the scattering, the photon energy is ϵ_1. The energy of the photon in the lepton rest frame is given by the Lorentz transformation $\epsilon_0' = \epsilon_0\gamma_0(1-\beta_0\mu_0)$, where $\mu_0 = \cos\theta_0$ and $\beta_0 = (1-\gamma_0^{-2})^{1/2}$. The energy of the scattered photon in the lepton rest frame is

$$\epsilon_1' = \frac{\epsilon_0'}{1+\dfrac{\epsilon_0'}{m_{\rm e}c^2}\,(1-\cos\Theta')}, \tag{11.2}$$

where $\cos\Theta' = \mu_0'\mu_1' + \sqrt{1-\mu_0'^2}\sqrt{1-\mu_1'^2}\cos(\varphi_1'-\varphi_0')$, Θ' being the angle between the incoming and the scattered photon directions in the lepton rest frame. We have $\mu_0' \approx -1$ by virtue of relativistic beaming, so we can approximate $\cos\Theta' \approx -\mu_1'$. The energy of the scattered photon in the lepton rest frame ϵ_1' is determined using the full inverse Compton differential cross section from QED (see Sec. 11.2.3). Given ϵ_1', the scattering angle in this frame is deduced using Eq. (11.2):

$$\mu_1' = \frac{m_e c^2}{\epsilon_1'} - \frac{m_e c^2}{\epsilon_0'} - 1. \tag{11.3}$$

Finally, another Lorentz transformation gives the energy of the scattered photon back in the FIDO frame:

$$\epsilon_1 = \gamma_0(1+\beta_0\mu_1')\epsilon_1'. \tag{11.4}$$

Thus, once the angle of the incoming photon θ_0 is randomly drawn, we only need to draw the energy of the scattered photon in the lepton rest frame ϵ_1'. Let us summarize the Monte-Carlo scheme for IC scattering.

1. Compute the FIDO-measured Lorentz factor $\gamma_0 = \sqrt{1+h^{jk}u_j u_k}$.
2. Compute $\delta\tau$ and determine whether the lepton scatters a soft photon or not.
3. Draw uniformly the random variable $\mu_0 \in [-1,1]$ (because the radiation field is assumed to be isotropic).

4. Compute $\epsilon_0' = \epsilon_0\gamma_0(1 - \beta_0\mu_0)$.
5. Draw the scattered photon energy ϵ_1' (see Sec. 11.2.3).
6. Compute μ_1' using Eq. (11.3).
7. Deduce the energy of the scattered photon in the FIDO frame using Eq. (11.4).
8. Update the energy of the incident lepton to $\gamma_0 m_e c^2 + \epsilon_0 - \epsilon_1$, by the conservation of energy.

In the code, we create a high-energy photon at the location of the scattering lepton, with energy ϵ_1. Assuming strong relativistic beaming again, the direction of the scattered photon in the FIDO frame is the same as that of the scattering lepton.

11.2.2 Drawing a random variable from a probability distribution function

The essence of a Monte-Carlo method is to sample quantities from their normalized probability distribution functions (PDF). Let us sample a quantity $X \in [a, b]$, with PDF P defined on $[a, b]$, and define the cumulative distribution function (CDF) as

$$\mathcal{P}(x) = \int_a^x P(X)\,\mathrm{d}X. \tag{11.5}$$

$\mathcal{P}$ is monotonic, strictly increasing, and ranges from 0 to 1 for $x \in [a, b]$. $\mathcal{P}(x)$ represents the probability that the random variable X should lie in $[a, x]$.

Let us state an important result in the context of Monte-Carlo algorithms. If ξ is a random variable uniformly distributed in $[0, 1]$, then there exists $x \in [a, b]$ such that $\xi = \mathcal{P}(x)$. The fundamental principle states that x is distributed from the PDF P (Kalos & Whitlock, 2008).

Proof

Let ξ be a uniformly distributed variable on $[0, 1]$. Because ξ is uniform, for any real $y \in [0, 1]$, the probability that $\xi < y$ is y, which we write $\mathbb{P}[\xi < y] = y$. Applying this equality with $y = \mathcal{P}(x)$, and using the fact that $\mathcal{P}$ and its inverse $\mathcal{P}^{-1}$ are monotonic and increasing, we find that $\mathcal{P}(x) = \mathbb{P}[\xi < \mathcal{P}(x)] = \mathbb{P}\left[\mathcal{P}^{-1}(\xi) < x\right]$. By definition, this means that $\mathcal{P}$ is the CDF of the random variable $\mathcal{P}^{-1}(\xi)$. As a result, $x = \mathcal{P}^{-1}(\xi)$ is distributed from the PDF P.

This method is used to draw the initial angular particle distribution, or the energies of the particles if the particle population is chosen to be thermal. Sampling from a given PDF requires to invert the equation $\xi = \mathcal{P}(x)$. This can be done by interpolation, if a table of $\mathcal{P}$ has already been computed.

11.2.3 Drawing the energy

So far we have only written down the "kinematic" steps of the inverse Compton algorithm, overlooking the probabilistic part of the interaction. Probabilities come into play when determining the energy of the scattered photon ϵ_1'. By definition of the differential cross section, the number of scattered photons with an energy $\in [\epsilon_1', \epsilon_1' + \mathrm{d}\epsilon_1']$ and within a solid angle $\in [\Omega_1', \Omega_1' + \mathrm{d}\Omega_1']$ in the lepton rest frame is

$$\frac{\mathrm{d}N'}{\mathrm{d}t'\,\mathrm{d}\epsilon_1'\,\mathrm{d}\Omega_1'} = cn_0\frac{\mathrm{d}\sigma_{\mathrm{IC}}}{\mathrm{d}\epsilon_1'\,\mathrm{d}\Omega_1'}. \tag{11.6}$$

The probability that a scattered photon has its energy $\in [\epsilon_1', \epsilon_1' + \mathrm{d}\epsilon_1']$ is

$$P\left(\epsilon_1'\right) = \frac{\int \mathrm{d}\Omega_1' \, \mathrm{d}\sigma_{\mathrm{IC}}/\,\mathrm{d}\epsilon_1' \, \mathrm{d}\Omega_1'}{\int \mathrm{d}\epsilon_1' \, \mathrm{d}\Omega_1' \, \mathrm{d}\sigma_{\mathrm{IC}}/\,\mathrm{d}\epsilon_1' \, \mathrm{d}\Omega_1'} = \frac{\int \mathrm{d}\Omega_1' \, \mathrm{d}\sigma_{\mathrm{IC}}/\,\mathrm{d}\epsilon_1' \, \mathrm{d}\Omega_1'}{\sigma_{\mathrm{IC}}}. \tag{11.7}$$

We need to sample the energy ϵ_1' from this PDF, at fixed ϵ_0'. To do this, we apply the method described in the previous section, and compute the CDF. Let us define the parameter $y = \epsilon_1'/\epsilon_0'$. Integrating the differential cross section over the angles will remove the Dirac function

$$\delta\left(\epsilon_1' - \frac{\epsilon_0'}{1 + \dfrac{\epsilon_0'}{m_{\mathrm{e}}c^2}\left(1 - \cos\Theta'\right)}\right) = \frac{m_{\mathrm{e}}c^2}{{\epsilon_1'}^2}\delta\left(\cos\Theta' - 1 - \frac{m_{\mathrm{e}}c^2}{\epsilon_0'} + \frac{m_{\mathrm{e}}c^2}{\epsilon_1'}\right). \tag{11.8}$$

The differential cross section (Eq. (8.9)) is rearranged as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_1' \, \mathrm{d}\epsilon_1'} = \frac{3\sigma_{\mathrm{T}}}{16\pi}\frac{m_{\mathrm{e}}c^2}{{\epsilon_0'}^2}\left(y + \frac{1}{y} - 1 + \cos^2\Theta'\right)\delta\left(\cos\Theta' - 1 - \frac{m_{\mathrm{e}}c^2}{\epsilon_0'}\left(1 + \frac{1}{y}\right)\right), \tag{11.9}$$

so that the PDF P reads (up to a proportionality factor), after integration:

$$P(y) \propto y + \frac{1}{y} + \frac{1}{\tilde{\epsilon}_0^2}\left(1 - \frac{2}{y} + \frac{1}{y^2}\right) + \frac{2}{\tilde{\epsilon}_0}\left(1 - \frac{1}{y}\right). \tag{11.10}$$

For conciseness, we have defined $\tilde{\epsilon}_0 = \epsilon_0'/m_{\mathrm{e}}c^2$ (which was noted as x in Sec. 8.1). Discarding the proportionality factor, a primitive of f is

$$F(y) = -\frac{1}{y\tilde{\epsilon}_0^2} + y\left(\frac{1}{\tilde{\epsilon}_0^2} + \frac{2}{\tilde{\epsilon}_0}\right) + \frac{y^2}{2} + \ln(y)\left(1 - \frac{2}{\tilde{\epsilon}_0} - \frac{2}{\tilde{\epsilon}_0^2}\right). \tag{11.11}$$

Before computing $\mathcal{P}_{\mathrm{IC}}$, we also need to determine the support of the PDF. The range of permitted y is determined by $-1 \leq \cos\Theta' \leq 1$, which yields

$$\frac{1}{1 + 2\tilde{\epsilon}_0} \leq y \leq 1. \tag{11.12}$$

The CDF $\mathcal{P}_{\mathrm{IC}}$ can be expressed as

$$\mathcal{P}_{\mathrm{IC}}(y) = \frac{F(y) - F\left(1/\left(1 + 2\tilde{\epsilon}_0\right)\right)}{F(1) - F\left(1/\left(1 + 2\tilde{\epsilon}_0\right)\right)}. \tag{11.13}$$

It is shown in Fig. 11.1 for several values of ϵ_0'.

We need to invert Eq. (11.13) in order to express y as a function of $\mathcal{P}_{\mathrm{IC}}$. Unfortunately, this is impossible to do analytically. Because computing the energy is a very frequent operation in the code (it must be performed for a significant fraction of leptons at almost every time step), this step need to be done as efficiently as possible. For that reason, rather than to invert the equation numerically, we decide to approximate the CDF, so we can invert it analytically. In the limit $\epsilon_1' \ll \epsilon_0'$, the primitive F reduces to $F(y) \simeq \ln(y)$. In this limit, the CDF reads

$$\mathcal{P}_{\mathrm{IC}}(y) \simeq \frac{\ln(y) + \ln\left(1 + 2\tilde{\epsilon}_0\right)}{\ln\left(1 + 2\tilde{\epsilon}_0\right)}. \tag{11.14}$$

The approximated CDF is also shown in Fig. 11.1 (see the black lines). We see that this approximation improves as $\tilde{\epsilon}_0$ decreases, but it remains acceptable even for large $\tilde{\epsilon}_0$ and y. Eq. (11.14) has the benefit of being easily invertible. In order to sample the scattered particle

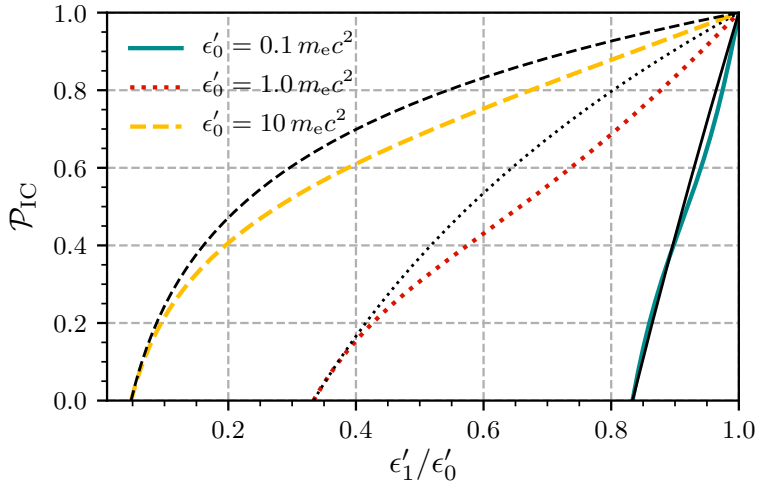


FIGURE 11.1: Analytical CDF for inverse Compton scattering, given by Eq. (11.13), for different values of ϵ'_0. The approximated CDF from Eq. (11.14) are also shown (black lines).

energy, we draw a random variable ξ uniformly distributed in $[0, 1]$, and obtain the scattered photon energy from

$$\epsilon'_1 = \epsilon'_0 \exp\left((\xi - 1)\ln\left(1 + \frac{2\epsilon'_0}{m_e c^2}\right)\right). \tag{11.15}$$

I have run sanity checks for the implementation of the algorithm in the code, by starting with a monoenergetic population of electrons and positrons with a FIDO-measured Lorentz factor γ_0, immersed into the soft radiation field. I compute the normalized spectrum of the high-energy photons after one time step, and plot it against the analytical solution given by Eq. (8.17), for different parameters Γ_ϵ. The result is shown in Fig. 11.2. The agreement is good in the Thomson regime and in the Klein-Nishina regime, so we are confident that all qualitative features of inverse Compton scattering are captured by the algorithm. The small discrepancies between the exact solution and our model are completely washed out when considering broader distributions of leptons in energy.

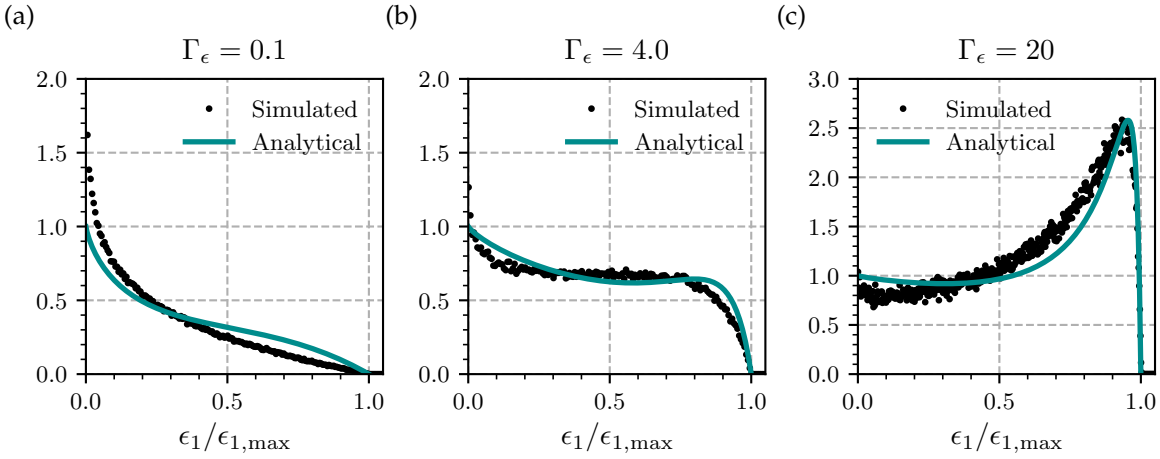


FIGURE 11.2: Simulated high-energy photon spectrum (black dots) for monoenergetic electrons and positrons propagating in a monoenergetic soft radiation field, for $\Gamma_\epsilon = 0.1$ (a), $\Gamma_\epsilon = 4$ (b) and $\Gamma_\epsilon = 20$ (c). The analytical prediction for the spectrum (Eq. (8.17)) is also shown as a blue solid line.

11.3 Pair production

11.3.1 Kinematics

We consider a high-energy photon of energy $\epsilon_1 \gg \epsilon_0$ in the FIDO frame, colliding with a soft photon with an angle θ_0. We use the same notations as in Sec. 8.2; quantities defined in the center-of-mass (CM) frame of the pair will be primed. The Monte-Carlo algorithm for pair production goes as follows.

1. Compute the FIDO-measured energy of a gamma photon $\varepsilon_1 = \sqrt{h^{jk} p_j p_k}$.
2. Draw uniformly the random variable $\mu_0 = \cos\theta \in [-1, 1]$ (because the radiation field is assumed to be isotropic).
3. Compute the Lorentz invariant $s = \epsilon_0 \epsilon_1 (1 - \mu_0)/2$. If $s \leq (m_e c^2)^2$, the photons are below the pair creation threshold: no pair is created and nothing happens.
4. Compute $\delta\tau$ and determine whether the photon interacts with a soft photon or not.
5. If an interaction event occurs, compute $\gamma_{\rm CM} = \epsilon_1/2\sqrt{s}$, $\beta_{\rm CM} = 1 - 2s{\epsilon_1}^2$, and the velocity of the pair in the CM frame β'.
6. Draw the angle $\mu' = \cos\theta'$ between the direction of the outgoing pair and the direction of motion of the center of mass, using the QED differential cross-section for pair creation (see Sec. 11.3.2).
7. Discard the high-energy photon from the simulation.
8. Compute the energies of the electron and the positron $E_\pm$ according to Eq. (8.32) and (8.33), and create a $e^\pm$ pair at the location of the high-energy photon with energies $E_\pm$.

We take the direction of propagation of the created pair to be along that of the primary gamma-ray. This approximation is valid provided $\gamma_{\rm CM} \gg 1$, which always holds since $\varepsilon_1 \gg \varepsilon_0$.

11.3.2 Drawing the angle

The probabilistic nature of the interaction is also involved when drawing the angle θ' between the outgoing pair and the center-of-mass momentum in the CM frame. According to Eq. (8.31), the probability that the pair makes an angle θ' such that $\cos\theta' \in [\mu', \mu' + {\rm d}\mu']$ at fixed β' is

$$P\left(\mu'\right) = \frac{{\rm d}\sigma_{\gamma\gamma}/\,{\rm d}\mu'}{\sigma_{\gamma\gamma}} = \frac{3}{16}\frac{\sigma_{\rm T}}{\sigma\gamma\gamma\left(\beta'\right)}\beta'\left(1-\beta'^2\right)\frac{1-\left(\beta'\mu'\right)^4+2\left(1-\beta'^2\right)\beta'^2\left(1-\mu'^2\right)}{\left(1-\left(\beta'\mu'\right)^2\right)^2}. \tag{11.16}$$

The CDF $\mathcal{P}_{\gamma\gamma}$ can be computed numerically from this PDF by integration and normalization; it is shown in Fig. 11.3 for various values of β'. We notice that the CDF has a behaviour very similar to the $x \mapsto \operatorname{arctanh}(x)$ function. Let us try to find a good approximation of $\mathcal{P}_{\gamma\gamma}$ which can be inverted analytically, and write

$$\mathcal{P}_{\gamma\gamma}\left(\mu'\right) \simeq p_1 + p_2 \operatorname{arctanh}\left(p_0 x\right), \tag{11.17}$$

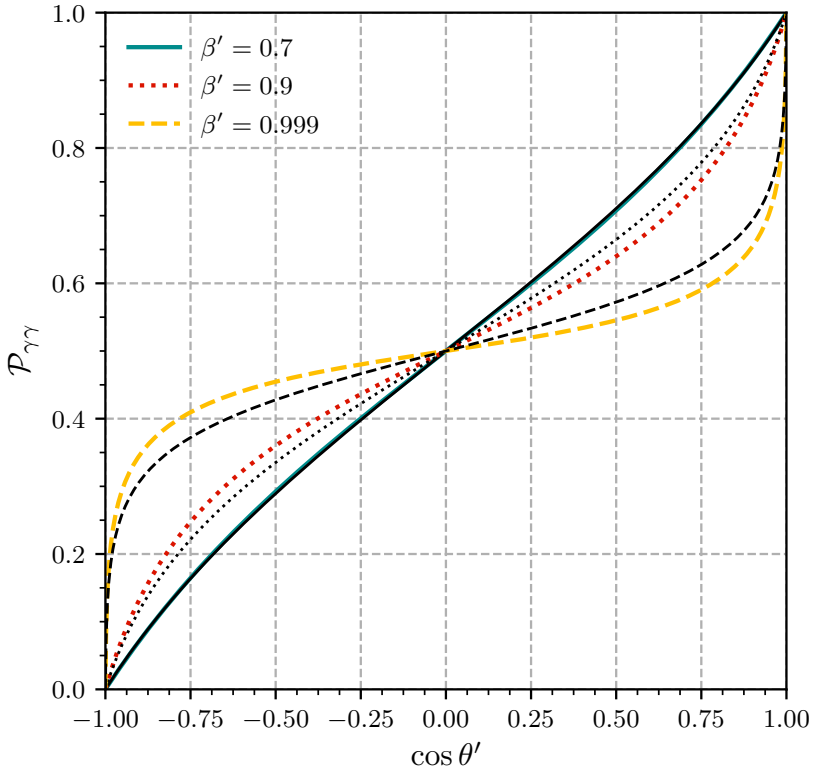


FIGURE 11.3: Analytical CDF for pair production given by Eq. (11.16), for different values of β'. The approximated CDF from Eq. (11.17) are also shown (black lines).

with three free parameters p_1, p_2 and p_0. We require that $\mathcal{P}_{\gamma\gamma}(0) = 1/2$ because the distribution is symmetrical with respect to $\mu' = 0$, and $\mathcal{P}_{\gamma\gamma}(1) = 1$ by definition of the CDF. This yields

$$\mathcal{P}_{\gamma\gamma}\left(\mu'\right) = \frac{1}{2}\left(1 + \frac{\operatorname{arctanh}\left(p_0 x\right)}{\operatorname{arctanh}\left(p_0\right)}\right). \tag{11.18}$$

The parameter p_0 determines the slope of $\mathcal{P}_{\gamma\gamma}$ at the origin $d = \mathcal{P}'_{\gamma\gamma}(0)$, which is $d = p_0/2\operatorname{arctanh}\left(p_0\right)$ in this approximation. Since the CDF is nothing but the derivative of the normalized PDF, the slope at the origin is

$$d = P(0) = \frac{3}{16}\frac{\sigma_{\mathrm{T}}}{\sigma\gamma\gamma\left(\beta'\right)}\beta'\left(1-\beta'^2\right)\left(1+2\beta'^2-2\beta'^4\right). \tag{11.19}$$

The exact slope at the origin, given by Eq. (11.19) is plotted in Fig. 11.4 against the approximated slope $p_0/2\operatorname{arctanh}\left(p_0\right)$ for the simple choice $p_0 = \beta'$. We notice that the two plots are quite similar, suggesting that the choice $p_0 = \beta'$ leads to a very decent approximation of the CDF by Eq. (11.17). All in all, we estimate the CDF for the angle μ' as

$$\mathcal{P}_{\gamma\gamma}\left(\mu'\right) \simeq \frac{1}{2}\left(1 + \frac{\operatorname{arctanh}\left(\beta' x\right)}{\operatorname{arctanh}\left(\beta'\right)}\right). \tag{11.20}$$

The approximated CDF is also shown in Fig. 11.3 (see the black lines): it reproduces well the features of the exact CDF. Here again, we have chosen simplicity over the accuracy. It is possible to check that computing the exact analytical slope (Eq. (11.19)) and inserting it into Eq. (11.17) leads to an estimate of the CDF almost indistinguishable from the exact one. However, we find that this computation is heavier, and the essential features of the pair production are modeled with our simpler choice of slope. Eq. (11.17) can be easily inverted

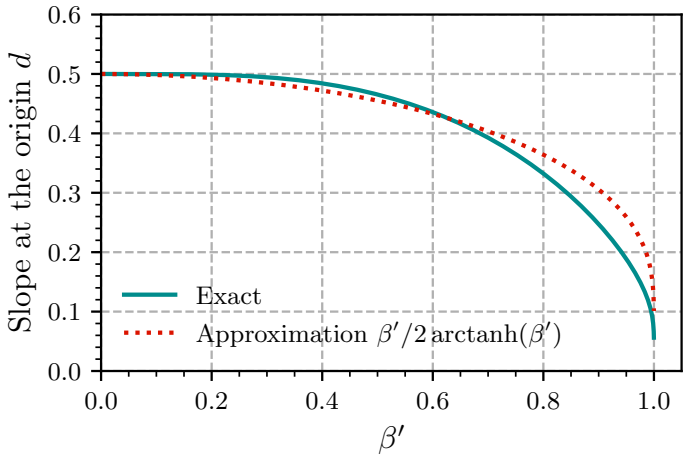


FIGURE 11.4: Exact slope at the origin of the pair production CDF (Eq. (11.16)), given by Eq. (11.19) (solid line). The approximation $d = \beta'/2\,\mathrm{arctanh}\,(\beta')$ is also shown as a dotted line.

analytically. In order to sample the outgoing angle of the pair, we draw a random variable ξ uniformly distributed in $[0, 1]$, and obtain the variable μ' from

$$\mu' = \frac{1}{\beta'} \tanh\left((2\xi - 1)\,\mathrm{arctanh}\,(\beta')\right). \tag{11.21}$$

I have also run sanity checks for the pair production algorithm. Starting with a monoenergetic population of high-energy photons, with a FIDO-measured energy ϵ_1, immersed into the soft isotropic and monoenergetic radiation field, we compute the energy spectrum of the electrons after one time step. The solution should match the analytical prediction given in Eq. (8.34). The result is shown in Fig. 11.5 for two values of the parameter ζ. The agreement is good both close to the pair creation threshold ($\zeta \simeq 1$), where the electron and the positron have the same energy in the FIDO frame, and far from the threshold ($\zeta \gg 1$), where the pair's energy is asymmetric.

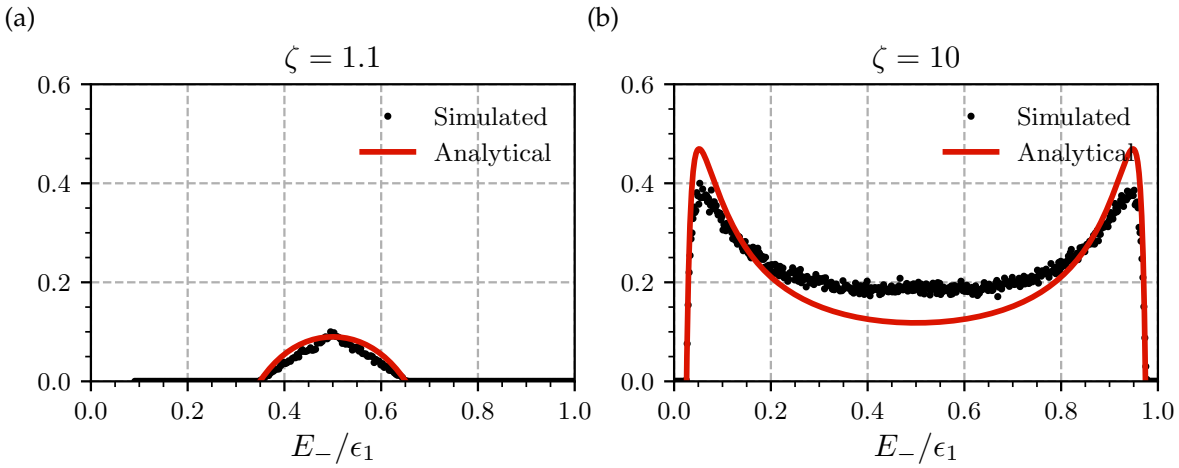


FIGURE 11.5: Simulated energy spectrum (black dots) of electrons produced by pair creation of monoenergetic photons propagating in a monoenergetic soft radiation field, for $\zeta = 1.1$ (a) and $\zeta = 10$ (b). The analytical prediction (Eq. (8.34)) is also shown as a red solid line.

Part IV

Results

"S'il n'y a pas de solution, c'est qu'il n'y a pas de problème."

"If there is no solution, then there is no problem."

JACQUES ROUXEL, Proverbe Shadok

Chapter 12

Simulations of pair discharges

In this section, I present 2D global GRPIC simulations of a black-hole magnetosphere carried out using the `GRZeltron` (see Chap. 10) code supplemented with the radiative transfer module (see Chap. 11). My aim is to study the spark gap dynamics by performing self-consistent global simulations of pair cascades in a black-hole magnetosphere, in the context of low-luminosity AGN, and to evaluate whether it is possible to fill the magnetosphere and activate the Blandford-Znajek process.

This chapter is partly adapted from my paper "Multidimensional Simulations of Ergospheric Pair Discharges around Black Holes", published in *Physical Review Letters* in 2020 (Crinquand et al., 2020). I presented these results at the "Simulating the evolution and emission of relativistic outflows" workshop in Meudon, in November 2019, and at the "Texas Symposium on Relativistic Astrophysics" in Portsmouth, in December 2019.

12.1 Preliminaries

12.1.1 Numerical setup

In this work, the black hole is endowed with a monopole magnetic field. The initial magnetic configuration is fully determined by the 4-potential A_μ, which is prescribed by Eq. (6.24). Although unphysical, this magnetic configuration has several upsides.

- The results can be directly compared to the Blandford-Znajek analytical solution, which assumes a magnetic monopole.
- The intrinsic physical properties of the spark gaps can be captured without interference from more complex structures, such as current sheets.
- A split monopole is a realistic model for the field lines penetrating the ergosphere on each hemisphere, irrespective of the magnetosphere's large-scale structure (Komissarov & McKinney, 2007; Parfrey et al., 2019). This is because field lines threading the event horizon are transported by the inflowing plasma inside the inner light surface.

The differences between simulating a full monopole and a split monopole mostly amount to the presence of an equatorial current sheet. Its impact will be explored in Chap. 13. The numerical setup is represented in Fig. 12.1. The location of the gap is represented as *a priori* unknown. The simulation domain is $r \in [r_{\rm min} = 0.9\, r_{\rm h}, r_{\rm max} = 16\, r_{\rm g}]$, $\theta \in [0, \pi]$. The inner boundary $r_{\rm min}$ is put well within the event horizon, so the boundary conditions for the fields at $r = r_{\rm min}$ do not impact the simulation. The spatial grid points are uniformly spaced in $\log(r)$ and θ, so the polar axes are well resolved, as well as the high-density regions close to the black hole. We mimic an open outer boundary using an absorbing layer from $r_{\rm abs} = 15\, r_{\rm g}$ to the end of the box at $r_{\rm max}$ (Cerutti et al., 2015). In this layer, a resistivity term is implemented in order to damp the electromagnetic waves arriving at $r_{\rm abs}$. For example, in this layer, Maxwell-Faraday's equation becomes

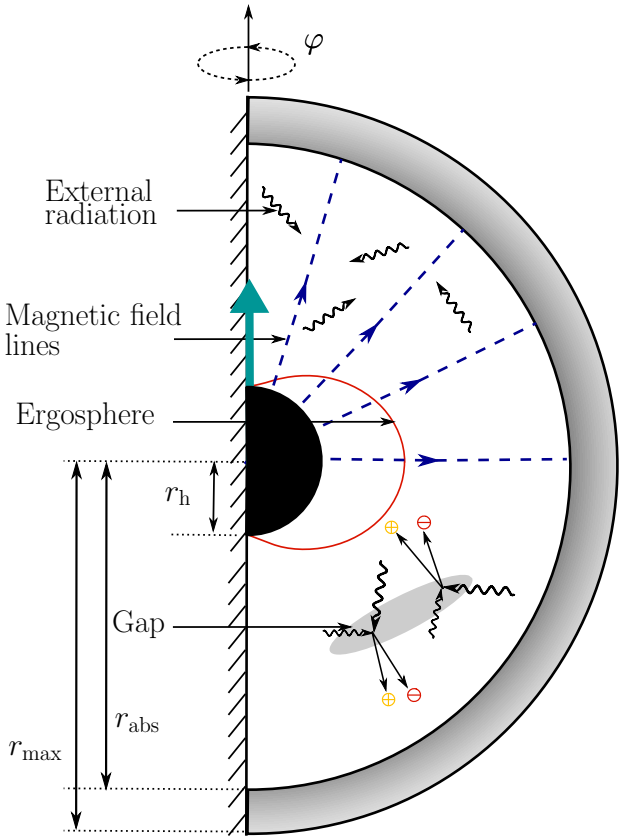


FIGURE 12.1: Numerical setup of the simulation. The shaded zone represents the absorbing layer, whereas the blue thick arrow represents the black hole spin.

$$\partial_t \boldsymbol{B} = -c\boldsymbol{\nabla} \times \boldsymbol{E} - \lambda\left(r\right)\boldsymbol{B}, \tag{12.1}$$

where the resistivity profile $\lambda\left(r\right) = \lambda_0\left(r - r_{\rm abs}\right)^3$ (for $r \geq r_{\rm abs}$) is chosen so as to avoid reflections of waves.

Particles are removed if $r \leq r_h$ or $r \geq r_{\rm abs}$. The runs have a grid resolution $2040\left(r\right) \times 2048\left(\theta\right)$, with the requirement that the plasma skin depth is resolved everywhere. This is tested *a posteriori*, since the plasma density is one of the unknowns. Unless specified otherwise, all our simulations were run with the spin parameter set at $a = 0.99$, in order to maximize the magnitude of the electric field, the spatial extent of the ergosphere and the magnitude of energy extraction.

12.1.2 Parameters

Electrons, positrons and gamma-ray photons interact with a background radiation field of soft photons. For simplicity, it is assumed that the radiation field is time-independent, uniform, isotropic and mono-energetic, with energy ϵ_0 and density n_0. There is no feedback of the simulation on this radiation field. The upscattered photons and created leptons are assumed to propagate along the same direction as their high-energy parents, reflecting strong relativistic beaming. The fiducial optical depth of both processes is τ_0 (see Chap. 8).

The normalized magnetic field is $\tilde{B}_0 = r_{\rm g}(eB_0/m_{\rm e}c^2)$, and the normalized energy of background photons is $\tilde{\epsilon}_0 = \epsilon_0/m_{\rm e}c^2$. $\tilde{B}_0$ is the ratio between the gravitational radius of

the black hole and the fiducial Larmor radius of particles. For a plasma multiplicity $\kappa \simeq 1$ and a close to 1, the fiducial plasma magnetization is $\sigma_0 = B_0^2/4\pi n_{\rm GJ} m_{\rm e} c^2 \simeq \tilde{B}_0$. Hence, the plasma is very magnetized, and flows approximately along magnetic field lines. The simulations start by having high-energy photons distributed uniformly and isotropically from $r = r_{\rm h}$ to $r = 4\,r_{\rm g}$, with the energy $\epsilon_1 \simeq m_{\rm e}c^2/\tilde{\epsilon}_0$. The photons quickly pair produce, "igniting" the cascade. From there, the system takes around 50 to 100 light-crossing times $r_{\rm g}/c$ to reach a steady state, as determined by the total number of particles in the box. To save computation time, photons with energy below the pair creation threshold are discarded from the simulation. Indeed, they are no longer able to participate in the dynamics.

Three dimensionless parameters define the physical conditions around the black hole: $\tilde{B}_0$, $\tilde{\epsilon}_0$ and τ_0. In M87*, the magnetic field at the event horizon is estimated to be $B_0 \simeq 10{-}100$G (Neronov & Aharonian, 2007; Event Horizon Telescope Collaboration et al., 2019b), which translates into $\tilde{B}_0 \simeq 10^{13-14}$. The soft background photon field peaks at $\epsilon_0 \simeq 1\,{\rm meV}$, so that $\tilde{\epsilon}_0 \simeq 10^{-9}$ (Prieto et al., 2016). The peak opacity for gamma-ray photons is at the energy $\epsilon \simeq 10^3$ TeV. The optical depth at the target energy 1 eV is uncertain, but is likely to be $\lesssim 10^3$ (Levinson & Rieger, 2011; Mościbrodzka et al., 2011). The typical density needed to screen the vacuum parallel electric field is the Goldreich-Julian number density $n_{\rm GJ} = B_0\omega_{\rm BH}/(4\pi ce)$ (see Eq. (6.36)), with $\omega_{\rm BH}$ the angular velocity of the black hole.

Let us denote as γ the typical Lorentz factor of a pair-producing lepton. In order for vacuum to break down, gamma-rays produced by inverse Compton scattering must verify $\tilde{\epsilon}_1\tilde{\epsilon}_0 \gtrsim 1$. This implies that inverse Compton scattering must occur in the Klein-Nishina regime. Indeed, this regime is characterized by a scattered energy $\epsilon_1 \simeq \gamma m_{\rm e}c^2$ and $\epsilon_0{}' \gtrsim m_{\rm e}c^2$. The latter equality implies $\epsilon_0 \gtrsim m_{\rm e}c^2/\gamma$, so that the condition for vacuum breakdown $\tilde{\epsilon}_1\tilde{\epsilon}_0 \gtrsim 1$ is satisfied. Photons scattered in the Thomson regime are not energetic enough to produce pairs. The typical behavior of an electron is shown in Fig. 12.2. In Fig. 12.2(a), the Lorentz factor shows steep drops at first, indicative of Klein-Nishina losses. When the electron has slowed down sufficiently, its cooling is more continuous, operating in the Thomson regime. As can be seen in Fig. 12.2(b), the electron was emitted in the ergoregion and accelerated during its outward flight.

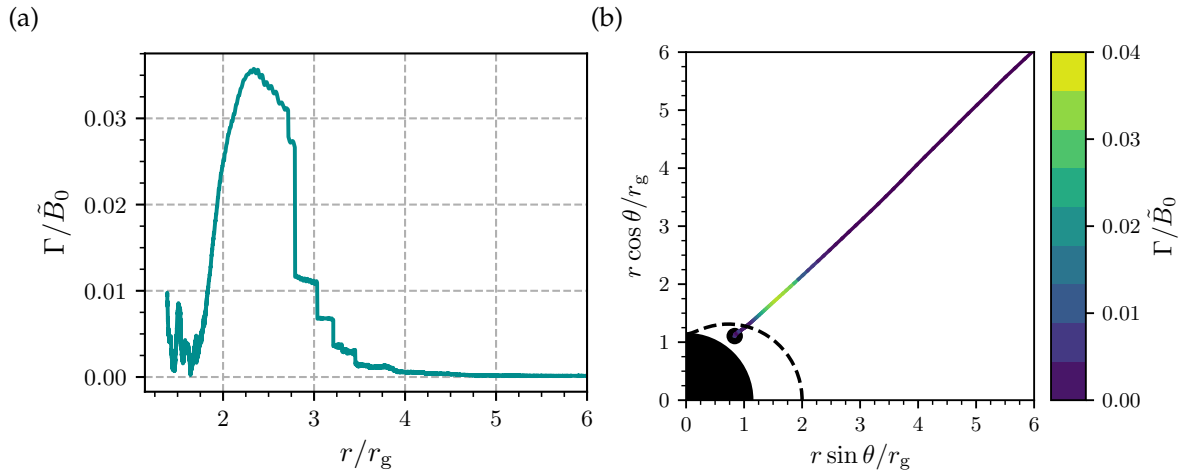


FIGURE 12.2: (a) FIDO-measured Lorentz factor Γ of a single electron, normalized by $\tilde{B}_0$, as a function of its distance to the black hole. (b) Trajectory of the same electron in the simulation box.

12.1.3 Downscaling

Let us assume that an electron is accelerated in a gap of size H by some unscreened electrostatic field. In a quasi force-free magnetosphere, one does not expect large unscreened electric fields parallel to the magnetic field. We can take B_0 as an upper limit on the parallel electric field, in the case of high spins $a \simeq 1$. We can also assume that the gap has macroscopic size, so that H is a significant fraction of r_g. Consequently, the maximum Lorentz factor γ_max that leptons can reach is of the order of $\tilde{B}_0$. We also define γ_s as the typical Lorentz factor of secondary particles that have been pair produced, and $\gamma_\mathrm{e} - 1$ the binding energy necessary for particles to escape the magnetosphere, with γ_e not much larger than 1. We focus our work on AGN characterized by the following separation of scales:

$$\gamma_\mathrm{e} \ll \gamma_s \ll \gamma_\mathrm{max}. \tag{12.2}$$

The first inequality states that the bulk of pair-produced particles should be energetic enough to from escape the black hole, whereas the second one indicates that leptons can scatter photons in the Klein-Nishina regime, above the pair creation threshold.

The cross-section of $\gamma\gamma$ pair production peaks near the pair creation threshold, so the bulk of pairs is created at $\gamma_\mathrm{s} \simeq 1/\tilde{\epsilon}_0$. The greater the ratio $\gamma_\mathrm{max}/\gamma_s \simeq \tilde{B}_0\tilde{\epsilon}_0$ gets, the higher the resulting plasma multiplicity κ will be (Timokhin & Harding, 2019). That is because an energetic lepton roughly loses an amount $\sim \gamma_s m_e c^2$ of energy every time a pair is created. This means that the parameter $\tilde{B}_0\tilde{\epsilon}_0$ must be chosen as large as possible to increase the multiplicity.

The typical cooling timescale of a lepton in the Thomson regime is given by $t_\mathrm{IC} \simeq r_\mathrm{g}/c\tau_0$. The fiducial optical depth τ_0 cannot be taken too high, otherwise the cooling would be so strong that particles could never reach the pair creation threshold and pair produce. However, this result is modified in the Klein-Nishina regime. We saw that the Klein-Nishina cross-section can be approximated by $\sigma_\mathrm{IC} \propto \sigma_\mathrm{T}/(\epsilon_0'/m_\mathrm{e}c^2)$. Evaluating the photon energy in the electron rest frame by $\epsilon_0' \simeq \gamma\epsilon_0$, we obtain a scaling for the cooling time t_KN in the Klein-Nishina regime:

$$t_\mathrm{KN} \simeq \frac{\gamma\tilde{\epsilon}_0}{\tau_0}\frac{r_\mathrm{g}}{c}. \tag{12.3}$$

If the optical depth is sufficently low, particle can leave the electrostatic gap without interacting with the soft photons.

It is impossible in practice to use realistic values for the parameters $\tilde{B}_0$ and $\tilde{\epsilon}_0$. Indeed, increasing the field strength implies increasing n_GJ and decreasing the skin depth $d_e = \sqrt{m_\mathrm{e}c^2/4\pi\kappa n_\mathrm{GJ}e^2}$, so that the resolution would need to go up. It is therefore challenging to find a part of the parameter space $(\tilde{B}_0, \tilde{\epsilon}_0, \tau_0)$ that is both accessible to simulations, and that has a large enough separation of scales to be relevant to the study of AGN. We have to reduce the separation of scales to computationally feasible ranges, while retaining the physical ordering given by Eq. (12.2). Altogether, we must choose $\tilde{\epsilon}_0$ as low as possible, but $\tilde{B}_0\tilde{\epsilon}_0$ large enough so the inequality $\tilde{B}_0\tilde{\epsilon}_0 \gg 1$ is satisfied. In practice we chose $\tilde{B}_0 = 5 \times 10^5$ and $\tilde{\epsilon}_0 = 5 \times 10^{-3}$. The product $\tilde{B}_0\tilde{\epsilon}_0 = 2500$ is still somewhat below a realistic value, which would be closer to 10^4 or 10^5, but we observed that this is sufficient to trigger time-dependent behavior and capture the physics of spark gaps. We ran four simulations for τ_0 going from 5 to 30. In this range of opacities, particles are not limited by radiative inverse Compton losses and can reach the pair creation threshold. Each simulation takes about 40 000 CPU hours.

12.2 Results

12.2.1 Structure of the simulation

We first focus on a fiducial simulation at $\tau_0 = 30$. The magnetic structure of this simulation is shown in Fig. 12.3. A negative toroidal magnetic field develops, while currents flow from the upper hemisphere ($\cos\theta > 0$) to the lower hemisphere ($\cos\theta < 0$). This is as it should be: these simulations have $\boldsymbol{\Omega} \cdot \boldsymbol{B} > 0$ in the upper hemisphere and $\boldsymbol{\Omega} \cdot \boldsymbol{B} < 0$ in the lower one, where $\boldsymbol{\Omega}$ is the black-hole angular velocity vector. In order to screen the electric field, the black-hole magnetosphere requires a negative poloidal current in the upper hemisphere and a positive current in the lower hemisphere.

The first question to address is whether the magnetosphere has reached a force-free state. To evaluate this, we verify some predictions made in Chap. 7 concerning force-free magnetospheres with a monopolar magnetic field. In the limit $|a| \ll 1$, both Eq. (7.10) and Eq. (7.11) should be valid. The toroidal magnetic field H_φ is a direct output of the code, whereas the angular velocity of the field lines is reconstructed using Eq. (6.47):

$$\Omega = -c\frac{E_\theta}{\sqrt{h}B^r}. \tag{12.4}$$

These quantities are shown in Fig. 12.4. The agreement is very good, although some discrepancies can be observed at the poles and in the equatorial plane. The discontinuities in Ω at the poles are likely to result from the boundary conditions for the fields. The slightly overestimated value of the force-free model for H_φ at $\theta = \pi/2$ could be due to a deficit in plasma

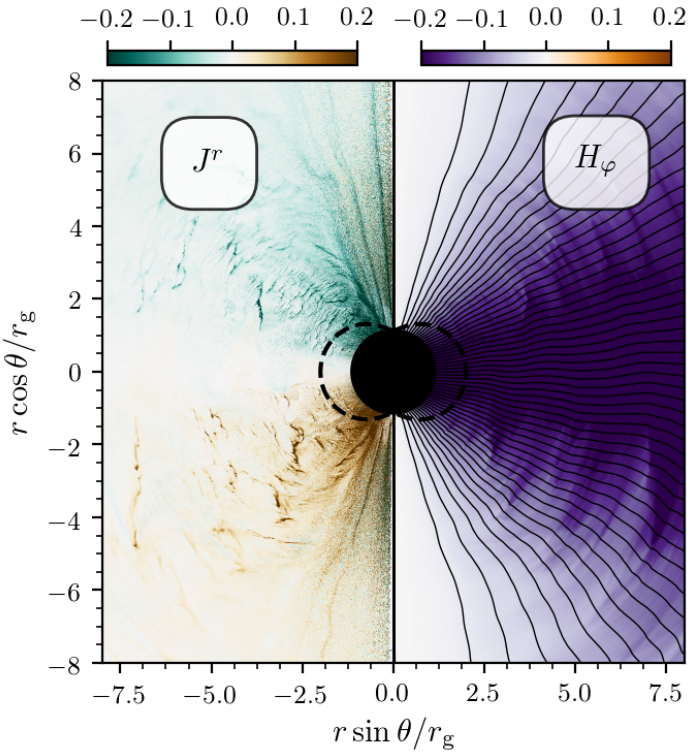


FIGURE 12.3: Snapshots of the normalized electric current density J^r/ecn_{GJ} (left panel) and normalized toroidal field $H_\varphi/B_0 r_\mathrm{g}$ (right panel) for the simulation at $\tau_0 = 30$. The black solid lines on the right panel are poloidal magnetic field lines. The thick dashed line indicates the ergosphere.

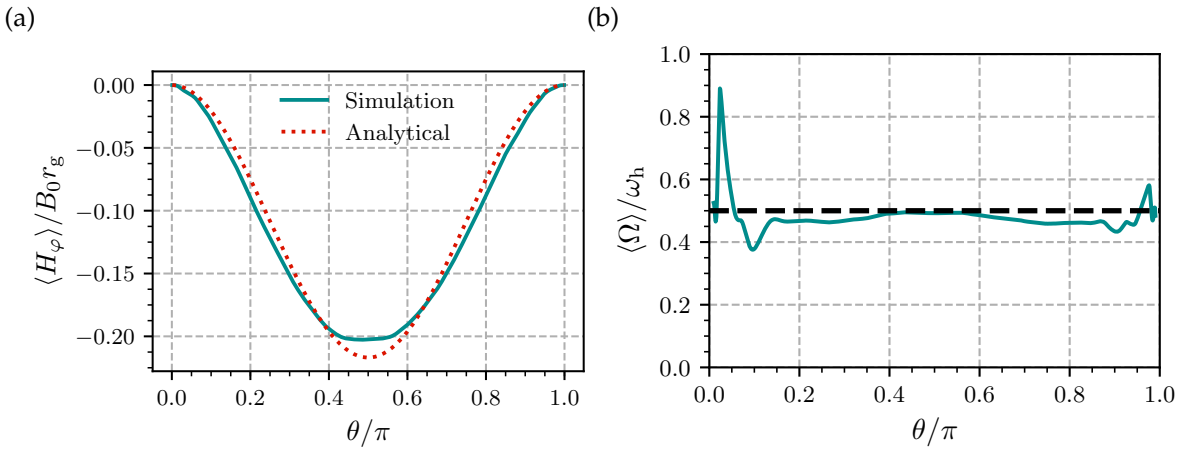


FIGURE 12.4: (a) Angular dependence of the normalized toroidal field $H_\varphi/B_0 r_g$ averaged over time and radius (solid line), and the analytical prediction from Eq. (7.11) (dashed line). (b) Angular dependence of the normalized angular velocity of the field lines Ω/ω_h averaged over time and radius. The dashed line indicates the analytical prediction from Eq. (7.10) $\Omega = \omega_h/2$ everywhere.

(see Fig. 12.5). Indeed, as we will see, the pair discharge mechanism is most efficient at intermediate latitudes. The angular velocity is practically constant over the simulation domain: all field lines are set into rotation at about $\omega_h/2$. Because both H_φ and Ω are nonzero, energy can be extracted from the black hole, as we will see in Sec. 12.2.4.

The electron and positron densities are shown in Fig. 12.5. Particles flow mainly radially,

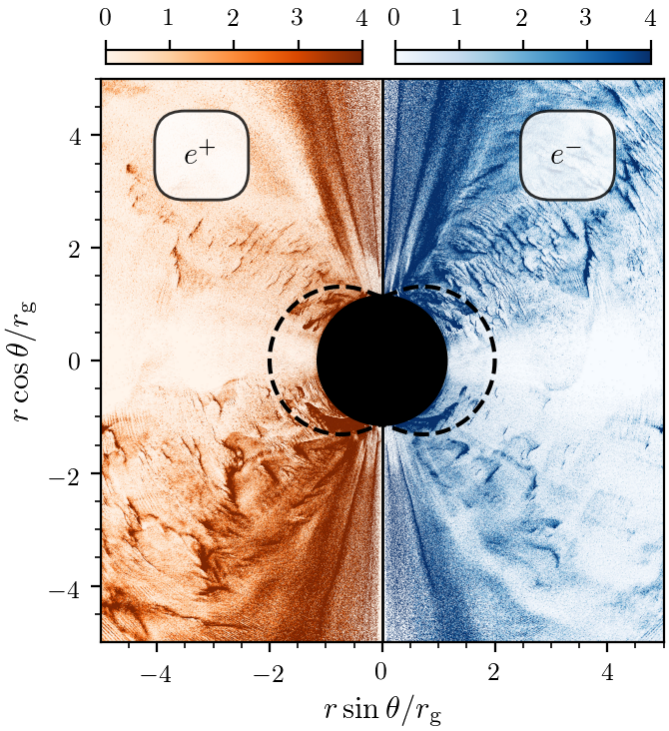


FIGURE 12.5: Snapshot of the steady-state normalized densities for positrons (left panel) and electrons (right panel), compensated by r^2. The thick dashed line indicates the ergosphere.

along the magnetic field lines. One species in the upper hemisphere has the same behavior as its anti-species in the lower hemisphere. Even if only electrons are required in the upper hemisphere and positrons in the lower hemisphere in order to carry the electric current and screen the electric field, both signs of charge are successfully ejected from the two hemispheres, filling the magnetosphere. It results that the electronic density is greater than the positron density for $z > 0$, and lower for $z < 0$. Still, the plasma remains quasi-neutral during the simulation: the plasma density is significantly larger than $\rho_{\rm GJ}/e$. A striking feature of Fig. 12.5 is that the plasma density is very clumpy, exhibiting shreds of pair plasma ejected intermittently. This is a smoking gun for variability in the pair creation process.

Fig. 12.6 shows the direction of motion of the electrons and positrons, measured using the grid-measured 3-velocity v^r, as well as the two light surfaces. The light surfaces are computed using the simulated angular velocity obtained with Eq. (12.4). Their shapes are consistent what was found in Sec. 4.1.5 in the case of a uniform angular velocity Ω. The size of the simulation box was set so as to include both light surfaces. This is crucial is order to get an outflowing solution.

Let us focus on the upper hemisphere. There is an electronic separation surface located exactly at the inner light surface. This is in line with the analysis from Sec. 7.2.4: below the inner light surface, all particles flow toward the black hole, whereas electrons above this surface flow out. There is no such clear distinction for positrons. As we will see, the direction of motion of the positrons depends on the sign of the parallel electric field $\boldsymbol{D}\cdot\boldsymbol{B}$. In any case, all particles within the inner light surface fall back to the black hole, whereas all particles beyond the outer light surface flow out. The situation is symmetric (switching positrons and electrons) in the lower hemisphere. The high-opacity simulations present similarities with the low plasma supply simulation in Parfrey et al. (2019), in particular regarding the role of the light surface. However, in our simulations all particles fly away from the black hole

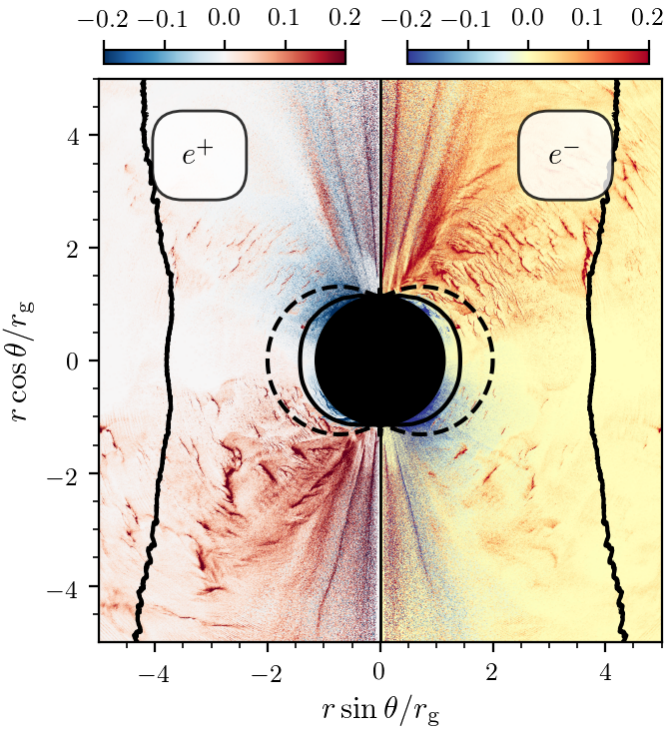


FIGURE 12.6: Snapshot of the steady-state normalized radial momentum flux $nv^r/n_{\rm GJ}c$ for positrons (left panel) and electrons (right panel). The thick dashed line indicates the ergosphere. The solid lines show the two time-averaged light surfaces.

outside of the outer light surface, as a consequence of the different magnetic configuration used. We ran a simulation at $\tau_0 = 30$ but with spin $a = 0.75$ and confirmed that the inner light surface retains the same role.

12.2.2 Dynamics of the gap

To better understand the dynamics of the gap, we plot several snapshots of the radial phase space (r, v^r) for electrons and positrons in Fig. 12.7, to illustrate how a pair creation burst develops. This phase space samples particles near $\theta_0 = \pi/4$ in the upper hemisphere. At this latitude, electrons are preferentially extracted from the magnetosphere to conduct the required current.

An electrostatic gap, characterized by a non-vanishing parallel electric field (as measured by FIDOs), opens *exactly* at the inner light surface. The typical normalized value of the unscreened electric field $\boldsymbol{D}\cdot\boldsymbol{B}/B^2$ ranges between 10^{-3} and 10^{-2}, which is similar to the *ad hoc* values used in Parfrey et al. (2019). This triggers particle acceleration, first in opposite directions, producing in turn high-energy photons that soon pair produce high-energy particles. As these secondary particles are created, they create a strong current which gradually screens the electric field parallel to the field lines. This screening process generates fluctuations of the parallel electric field. Importantly, the parallel electric field displays a positive overshoot in the course of the screening, which results in the outward acceleration of positrons too. This behavior had been derived by Levinson et al. (2005) in their study of pair creation oscillations, and was also seen by Philippov et al. (2020) in pulsar polar cap discharges. The intensity and duration of the bursts are highly variable. They have a spatial extent of a fraction of r_g (see Fig. 12.7), which is promising in the prospect of interpreting ultra-fast variability of AGN. At such high optical depth, the gap is narrow and very intermittent.

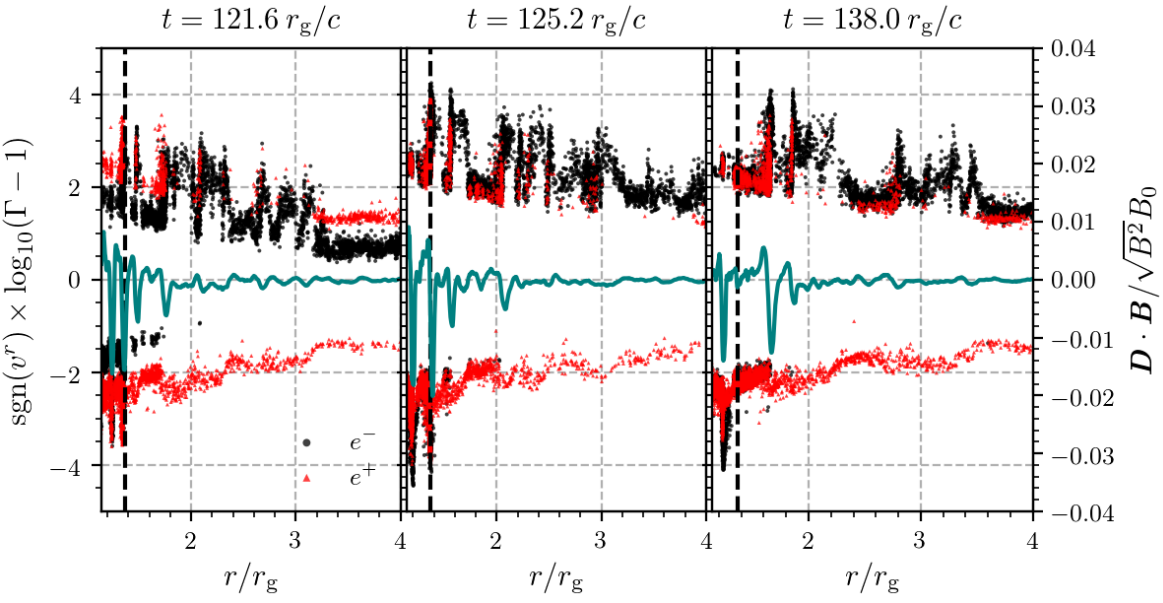


FIGURE 12.7: Snapshots of a phase space plot for electrons (black dots) and positrons (red triangles) sampled at $\theta_0 = \pi/4 \pm 0.02$ during a burst, for $\tau_0 = 30$. Particles are denoted by $\mathrm{sgn}(v^r)\log_{10}(\Gamma - 1)$, where Γ is its FIDO-measured Lorentz factor and v^r its radial 3-velocity. The blue solid line is the normalized unscreened electric field profile at θ_0. The vertical dashed line indicates the location of the light surface at θ_0. For clarity, only 20% of the particles are displayed.

As the gap opens, a burst of unscreened electric field either plunges inside the hole or moves outward. Subsequent pair creation occurs in this burst as it propagates, populating the magnetosphere with pair plasma. This is even more visible in the upper panel in Fig. 12.8, which shows a spacetime diagram and a 2D map of the pair creation rate. This highlights the variability of the gap as well as its small spatial extent. The inner light surface is the separation surface for the created pairs. Pair creation mainly takes place in the close magnetosphere, although it can extend beyond the ergosphere, as the screening of the outflowing troughs in the parallel electric field goes on. We note that the period of gap openings at the inner light surface is much smaller than the light-crossing time of the simulation box, so this is unlikely to be a numerical effect.

In this fiducial simulation, pairs are created at low altitudes so the gap can be screened efficiently. After a burst of pair creation, both electrons and positrons are expelled. The multiplicity κ of the plasma flow is high in the gap (around 10), and reaches 2 outside of a burst. We were unable to reproduce the very large multiplicities ($\simeq 10^3$) expected on theoretical grounds by Levinson & Rieger (2011). While this could be an effect of the reduced separation of scales, we observed no significant trend in the multiplicity when varying the parameters $\tilde{B}_0$, $\tilde{\epsilon}_0$ and τ_0.

The null surface, where the general relativistic Goldreich-Julian charge density vanishes, had been proposed as as plausible gap position (Chen & Yuan, 2020; Kisaka et al., 2020).

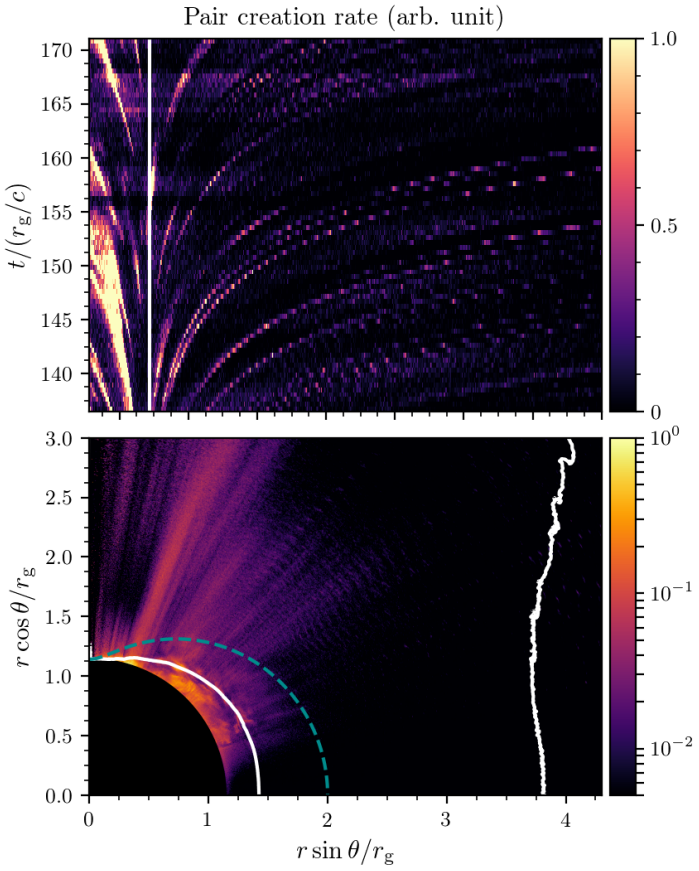


FIGURE 12.8: Top panel: spacetime diagram of the pair creation rate at $\theta_0 = \pi/4$ for $\tau_0 = 30$, in arbitrary units. The white solid line indicates the location of the inner light surface at $\theta = \pi/4$. Although pair creation is continuous in time in the simulations, trajectories look discretized because of downsampling. Bottom panel: 2D map of the time-averaged pair creation rate. The white solid line marks the time-averaged light surfaces, whereas the dashed line shows the ergosphere.

However, we find that it is irrelevant for the pair discharge, and that the inner light surface is where the gap forms. The specific location of the plasma supply in our simulations, which is imposed by the global structure of the magnetosphere, enforces the separation surface for electrons and positrons to lie close to the inner light surface. The effect of these surfaces can be most consistently taken into account in global 2D simulations.

12.2.3 Dependence on parameters

So far, we have focused on what could be called the high-opacity regime with $\tau_0 = 30$. The picture changes qualitatively as the parameter τ_0 is decreased, and we observe a transition between two regimes, illustrated in Fig. 12.9. The high-opacity solution is very close to being force-free, and the gaps opening at the inner light surface constitute a small correction to the structure of the magnetosphere. At low optical depths ($\tau_0 \lesssim 10$), pair formation occurs far from the black hole, resulting in a macroscopic low-density zone close to the horizon. The electric field remains unscreened in this zone, so a large and *steady* gap forms (see Fig. 12.9(a)). This causes large departure from the force-free behavior. Particles experience a large fraction of the full vacuum potential, which puts them deep into the Klein-Nishina regime. This results in a drop in the scattering cross section, pushing inverse Compton emission, and hence pair production, even further outward, outside of the acceleration zone. As τ_0 decreases, the location where the electric field is finally screened is pushed further and further from the black hole. This pushes the separation surface of positrons (in the upper hemisphere) further as well. In this regime, acceleration and pair creation are spatially decoupled. At even lower opacity the gap is so large and the particle energy so high that all particles escape the simulation before significantly pair producing. Intermediate opacity simulations display an intermediate regime: high-latitude field lines support a large and steady gap, whereas field lines close to the equator show the same time-dependent behavior as the high-opacity run.

We find that the gap size is controlled by the inverse Compton mean free path. At high opacity, the gap width is comparable to the mean free path in the Thomson regime r_g/τ_0.

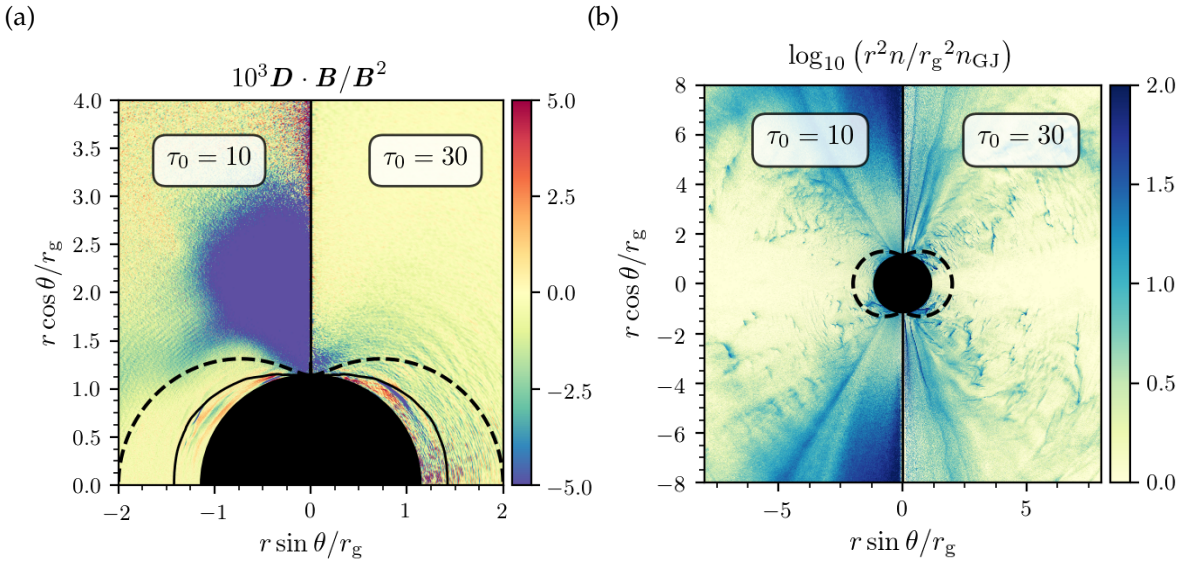


FIGURE 12.9: (a) Normalized parallel electric field $\boldsymbol{D}\cdot\boldsymbol{B}/\boldsymbol{B}^2$, rescaled by a factor 10^3, in the low-opacity case (left) and high-opacity case (right). The dashed line indicates the ergosphere and the solid line the inner light surface. (b) Logarithm of the normalized density compensated by r^2, in the low-opacity case (left) and high-opacity case (right).

The gap width H, measured using the unscreened electric field, is $H \simeq 0.06\, r_\mathrm{g}$ at $\tau_0 = 30$. At low opacity, the mean free path and the gap width become comparable to r_g. Particles reach high Lorentz factors in the gap, so the inverse Compton cross section drops, further increasing the mean free path. The optical depth is effectively lower than τ_0 close to the black hole.

It is also interesting to study a case where the inequality (12.2) is not well satisfied. We ran simulations with $\tau_0 = 30$, $\tilde{B}_0 = 10^3$ and $\tilde{\epsilon}_0 = 0.1$, an example of which is represented in Fig. 12.10. With this set of parameters, the inequality (12.2) reduces to $1 \ll 10 \ll 10^3$. In this case, there is no outflow of positrons in the upper hemisphere, no outflow of electrons in the lower hemisphere. All the created leptons are carrying the Goldreich-Julian current. The density is barely high enough to screen the electric field: the gap is never totally screened, and there are no wiggles in the electric field which can help to eject the other sign of charge. The magnetosphere does not reach a force-free state. This is because the produced pairs are not very energetic: barely enough particles make it into the deep Klein-Nishina regime and are able to pair produce significantly.

12.2.4 Energetics

The total Poynting flux of electromagnetic energy-at-infinity, which is the luminosity L of the black hole, can be extracted from the simulations according to Eq. (5.54):

$$L_\mathrm{EM}(r) = 2\pi \int_0^\pi \Pi^r\,(r,\theta)\,\sqrt{h}\,\mathrm{d}\theta. \tag{12.5}$$

The Poynting vector $\mathbf{\Pi}$ is given by Eq. (5.55), so that we have

$$\Pi^r = \frac{c}{4\pi\sqrt{h}}\left(E_\theta H_\varphi - E_\varphi H_\theta\right). \tag{12.6}$$

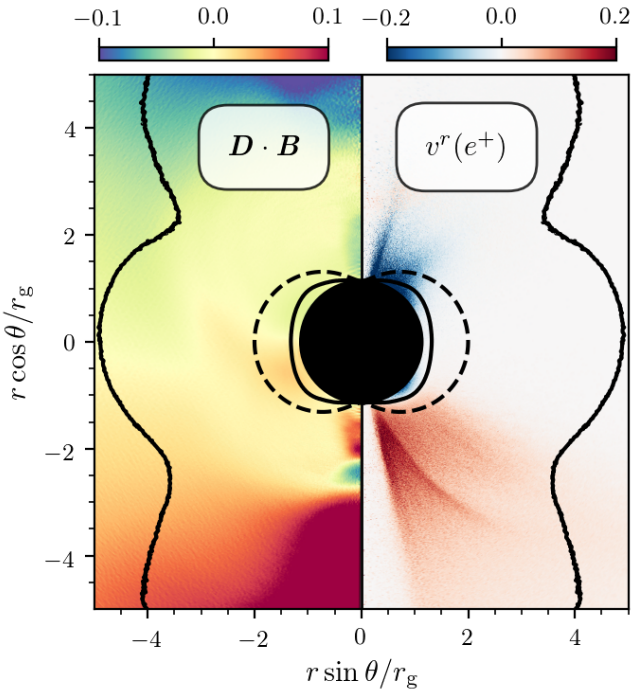


FIGURE 12.10: Snapshots of the steady-state normalized radial momentum flux $nv^r/n_\mathrm{GJ}c$ for positrons (right panel) and of the normalized parallel electric field $\boldsymbol{D}\cdot\boldsymbol{B}/\boldsymbol{B}^2$ (left panel) for a simulation at reduced separation of scales, with $\tau_0 = 30$, $\tilde{B}_0 = 10^3$ and $\tilde{\epsilon}_0 = 0.1$.

Fig. 12.11(a) shows the evolution of the electromagnetic luminosity (12.5) with the radius of the integration sphere. We find that the rate of energy extraction is consistent with the prediction $L_{\rm BZ}$ from the Blandford-Znajek mechanism, given by Eq. (7.18), at all opacities τ_0. This supports the role of the Blandford-Znajek mechanism in powering AGN jets and the possibility that inverse Compton scattering and photon-photon pair production processes can supply sufficient plasma to activate this mechanism.

We also observe that the luminosity decreases with increasing radius r, unlike what would happen in a perfectly force-free, ideal magnetosphere. This implies the presence of dissipation in the magnetosphere. The flux of energy-at-infinity of the particles in the simulation domain (electrons, positrons, and high-energy photons above the pair creation threshold) is shown in Fig. 12.11(b). At low opacity, a sizeable fraction of the Poynting flux (around 20%) is dissipated within the simulation domain, with about 1% being transferred to the simulated particles. At higher opacity, dissipation is smaller (around 5%) since the gap is narrow. The energy flux carried by particles is smaller by an order of magnitude. Dissipation increases with decreasing τ_0, because it scales approximately with the size of the gap. The bulk energy-at-infinity of the leptons within the ergosphere can be negative, but we find that they contribute little to the extraction of energy from the black hole in these simulations. This does not contradict the conclusions of Parfrey et al. (2019), because in their study most of the negative-energy electrons were located in an equatorial current sheet, while there is none in our simulations.

Note that the energy from photons below the pair creation threshold is lost to the simulation. Hence, most of the dissipated energy must rather be deposited in these particles. From Eq. (5.54), we can still estimate the power that they carry by integrating the dissipation rate $\boldsymbol{E}\cdot\boldsymbol{J}$ over the whole simulation box. At high optical depths, the dissipated power is around 3% of the output Poynting flux, approximately matching the decrease in the Poynting flux from $r_{\rm h}$ to $r_{\rm max}$. Hopefully, this energy can be tapped into in order to power VHE flares.

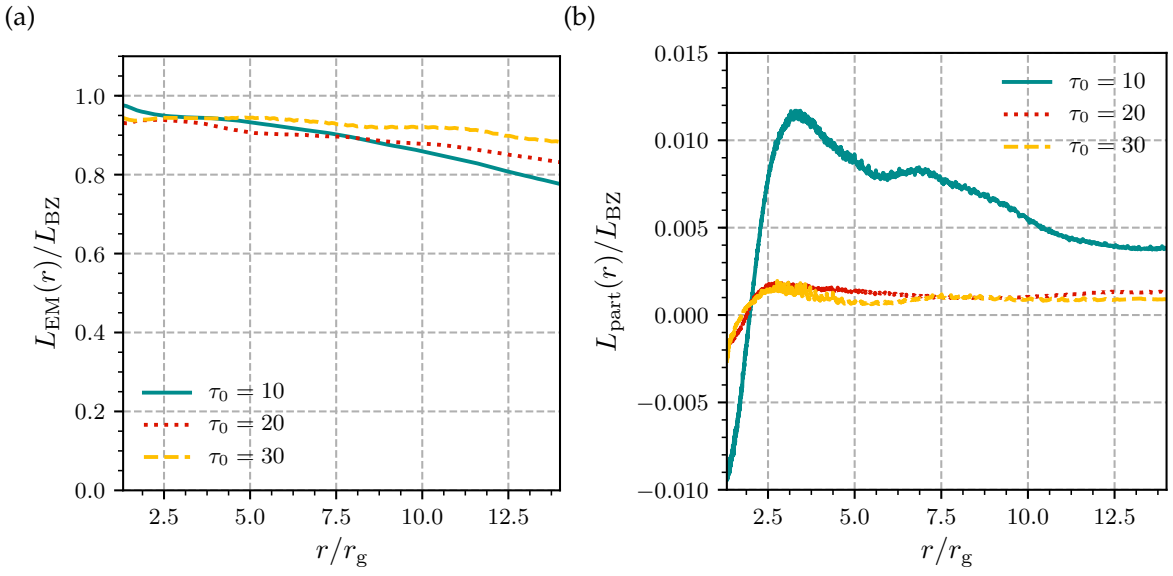


FIGURE 12.11: Total flux of electromagnetic energy-at-infinity $L_{\rm EM}$ (a) and particle flux of energy-at-infinity $L_{\rm part}$ (b) through a sphere of radius r, for $\tau_0 = 10$, 20 and 30. The particle flux comprises the contributions of electrons, positrons and high-energy photons above the pair creation threshold.

12.3 Discussion

Our results show some similarities with 1D models, but also important differences which justify the need for multi-dimensional simulations. Similarly to Chen & Yuan (2020), we find that the gap opens quasi-periodically. However, unlike them, we find that discharges happen at the inner light surface, whereas the null surface seems to play no role. Additionally, while their gap has a size $\gtrsim r_{\rm g}$, we find that the gap size is much smaller than the black hole size in the high-optical depth regime (although it remains much larger than the plasma skin depth). A major difference between 1D and 2D is that field lines do not all behave as a coherent entity. Therefore, the pair creation bursts have a smaller spatial extent and the time variability is higher in our simulations than in 1D models. On the other hand, we do not observe the quasi-steady, noisy state obtained by Levinson & Cerutti (2018), or by Chen & Yuan (2020) at low resolution. This might be because field lines can still weakly interact through the electric field in the (θ, φ) plane, retaining some coherence at small scale.

Chapter 13
Dynamics of the current sheet

In Chap. 12, I described the results of global multidimensional simulations of a black-hole magnetosphere, with a realistic treatment of radiative transfer. I was able to capture the intrinsic activity of spark gaps. However, as pointed out in Sec. 6.2, the magnetic structure was very simplistic. In the case of "isolated" magnetospheres, more realistic configurations with large-scale poloidal magnetic field should display an equatorial current sheet. This current sheet originates in the need to close the electric current system, since negative currents flow from both poles (if the spin axis is aligned with the magnetic field). It is unclear how such a current sheet can affect the pair discharge mechanism, and what are the relative contributions of the polar cap and the current sheet emissions.

It is still unclear what the geometry of the accretion flow and the magnetosphere looks like in the inner regions. At the very least, a situation with such a long-lived equatorial current sheet could come up if the accretion flow was truncated at large radius, so that accretion paused for a while. This is what happens in magnetically arrested disk simulations (Narayan et al., 2003). Still, large-scale and intermittent ergospheric current sheets are expected to develop naturally in accreting black-hole magnetospheres as well (Ripperda et al., 2020), highlighting the need to understand their importance. In this chapter, I explore more realistic magnetic configurations, no longer neglecting the equatorial reconnection activity.

This chapter is adapted from my paper "Synthetic gamma-ray light curves of Kerr black hole magnetospheric activity from particle-in-cell simulations", published in *Astronomy & Astrophysics* (Crinquand et al., 2021). I presented these results at the "Understanding the Most Energetic Cosmic Accelerators: Advances in Theory and Simulation" online workshop, organized in October 2020 by the PCTS at Princeton.

13.1 Numerical setup

13.1.1 Magnetic configuration

We saw that a large-scale poloidal magnetic configuration is favored by observations of the EHT and GRAVITY. Besides, it is the natural configuration to study the development of the Blandford-Znajek process and the launching of a jet. In this work, we will simulate a generic magnetic configuration with large-scale poloidal field.

> By contrast, Parfrey et al. (2015) and Mahlmann et al. (2020) have shown using general relativistic force-free simulations that this is not a necessary condition for efficient extraction of the black hole's energy. They were able to produce jets by the accretion of magnetized loops from the accretion disk, with zero net magnetic flux.

General relativistic magnetohydronamics (GRMHD) simulations of accretion flows also hint toward a paraboloidal geometry of the magnetic field lines (Komissarov & McKinney, 2007; McKinney et al., 2012). The initial poloidal magnetic field in the magnetosphere is

defined for $\theta \leq \pi/2$ by the following flux function (Tchekhovskoy et al., 2010)

$$A_\varphi(r,\theta) = B_0 r_\mathrm{g}^2 \left(\frac{r+r_0}{r_\mathrm{h}+r_0}\right)^\nu (1-\cos\theta), \tag{13.1}$$

where r_0 and ν are free parameters. The field lines have the opposite polarity in the lower hemisphere. We chose $r_0 = 10\,r_\mathrm{g}$ and $\nu = 3$ in our runs. The geometry of the initial magnetic field lines is shown in Fig. 13.1(a). It allows a current sheet to develop in the equatorial plane, without constraining its length *a priori*.

13.1.2 Conducting disk

A magnetosphere with such an initial magnetic field quickly dies out after a few tens of r_g/c. This lies in the fact that the current sheet extends to the outer boundary of the box, which is endowed with open boundary conditions. Magnetic reconnection at the current sheet ejects plasmoids and magnetic flux from the simulation box. Too much energy and magnetic flux are lost by the simulation box, so that the black hole almost completely expels the magnetic field lines threading it. Unlike pulsars for instance, black holes do not have a conducting surface and cannot sustain a magnetic field on their own. It is challenging to design a setup allowing the study of a non-vanishing current sheet in this context.

Therefore, I implemented a static and perfectly conducting disk as a boundary condition for the electromagnetic fields in the equatorial plane. This disk extends from its inner radius $r = r_\mathrm{in}$ to the outer boundary of the box, for $\theta \in [\pi/2 - \theta_0, \pi/2 + \theta_0]$, and we fixed $r_\mathrm{in} = 6\,r_\mathrm{g}$ and $\theta_0 = 0.02$ in all simulations. The resulting setup is represented in Fig. 13.1(b). This amounts to the condition of infinite conductivity given by Eq. (6.47), which was derived earlier, with $\boldsymbol{\Omega} = \boldsymbol{0}$ in the case of a non-rotating disk. Once $\boldsymbol{E}$ is prescribed by the infinite

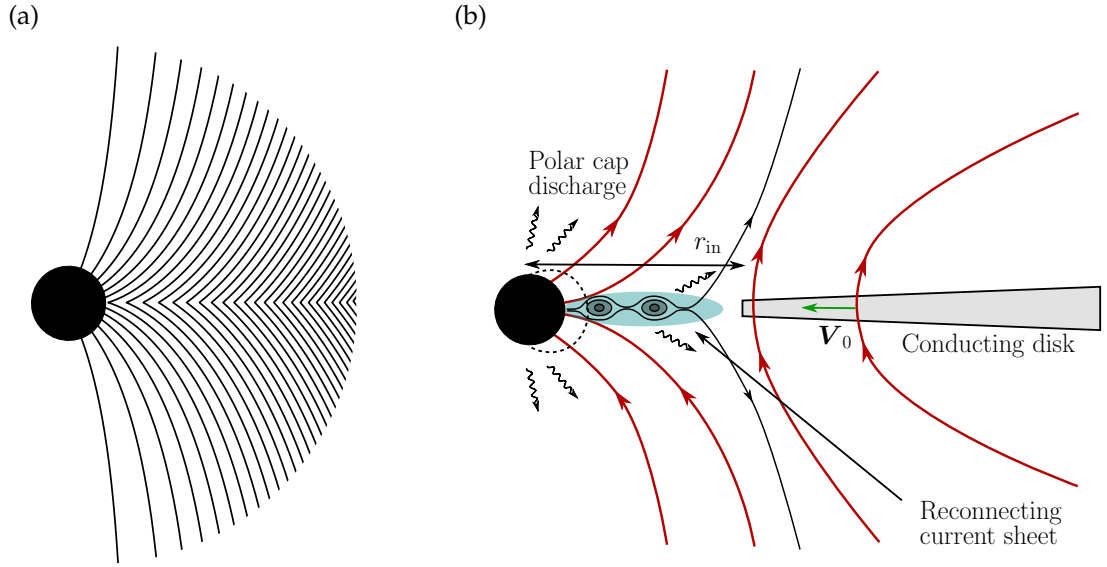


FIGURE 13.1: (a) Initial poloidal magnetic field lines, according to Eq. (13.1). (b) Schematic steady-state configuration of the simulations. Poloidal magnetic field lines are shown in red solid lines, except the last closed magnetic field line (in black). The equatorial current sheet (blue shaded area) is prone to the plasmoid instability. The conducting disk, represented by the shaded trapezoidal shape, extends from r_in to the outer edge of the simulation box. Two emitting zones are highlighted: the polar cap (low inclination with respect to the spin axis) and the current sheet.

conductivity condition, $\boldsymbol{D}$ is obtained from Eq. (5.41a). Denoting the initial magnetic field as $\boldsymbol{B}_{\rm in}$, the following conditions are enforced inside the disk:

$$E_r = E_\theta = E_\varphi = 0; \tag{13.2}$$

$$D_r = \frac{1}{\alpha} E_r = 0; \tag{13.3}$$

$$D_\theta = \frac{1}{\alpha}\left(E_\theta + \sqrt{h}\beta^r B^\varphi\right) = \frac{\beta^r}{\alpha}\sqrt{h}B^\varphi_{\rm in}; \tag{13.4}$$

$$D_\varphi = \frac{1}{\alpha}\left(E_\varphi - \sqrt{h}\beta^r B^\theta\right) = -\frac{\beta^r}{\alpha}\sqrt{h}B^\theta_{\rm in}; \tag{13.5}$$

$$B^\theta = B^\theta_{\rm in}; \tag{13.6}$$

$$H_\theta = \alpha h_{\theta\theta} B^\theta_{\rm in} + \sqrt{h}\beta^r D^\varphi. \tag{13.7}$$

Importantly, only the θ components of the magnetic fields $\boldsymbol{B}$ and $\boldsymbol{H}$ should be prescribed, otherwise the system is over-constrained and does not remain stable.

Magnetic field lines crossing this disk are frozen-in. With the addition of the conducting disk, the magnetic flux cannot escape the numerical domain, preventing our simulations from decaying entirely. Note that we exclude the study of the magnetic linkage that can exist between the black hole and the disk: no field lines connect the black hole to the disk. We do not claim to simulate a realistic accretion disk, but rather to provide physical conditions suitable to the study of the intrinsic behavior of the magnetosphere. The disk is merely included as a boundary condition for the fields: we are not interested in the zone surrounding the disk and focus on the magnetosphere itself, that is, the zone enclosed by the field lines crossing the ergosphere. For this reason, in all subsequent figures, we choose to leave the inner radius of the disk out of the represented domain. We also checked that there was no significant numerical diffusion, and hence no unphysical slippage of the field lines.

The numerical domain extends from $r_{\rm min} = 0.9\, r_{\rm h}$ to $r_{\rm max} = 10\, r_{\rm g}$. The spin parameter a is fixed at 0.99. Here, we focus on the high-optical depth regime, and run simulations with $\tau_0 = 30$, 50 and 80. We keep $\tilde{B}_0 = 5 \times 10^5$ and $\tilde{\epsilon}_0 = 5 \times 10^{-3}$ fixed in all simulations. Each simulation takes about $60\,000$ CPU hours.

13.2 Results

We first describe the general features of our simulations, before addressing the influence of magnetic field transport and the long-term evolution of the magnetosphere.

13.2.1 General features

The structure of the magnetosphere is shown in Fig. 13.2; it is consistent with the general picture that was drawn in Sec. 6.2.4. The right panel shows H_φ, which quantifies the poloidal current. This toroidal field is nonzero on the field lines connected to the black hole, penetrating the ergosphere; therefore a nonvanishing flux of energy and angular momentum can flow along those lines. The left panel shows the radial component of the current density. The electric current system is consistent with what is expected for a black-hole magnetosphere in the force-free regime with $a > 0$. Our simulations have $\boldsymbol{\Omega} \cdot \boldsymbol{B} > 0$ in both hemispheres. In the upper hemisphere, an electric field pointing toward the black hole is gravitationally induced by the frame-dragging of magnetic field lines. Negative poloidal currents are generated, which help screen the initial nonzero $\boldsymbol{D} \cdot \boldsymbol{B}$, thus giving rise to a negative H_φ. The situation is opposite in the lower hemisphere. By symmetry, H_φ must vanish in the equatorial plane. The subsequent current sheet carries a positive electric current, closing the electric

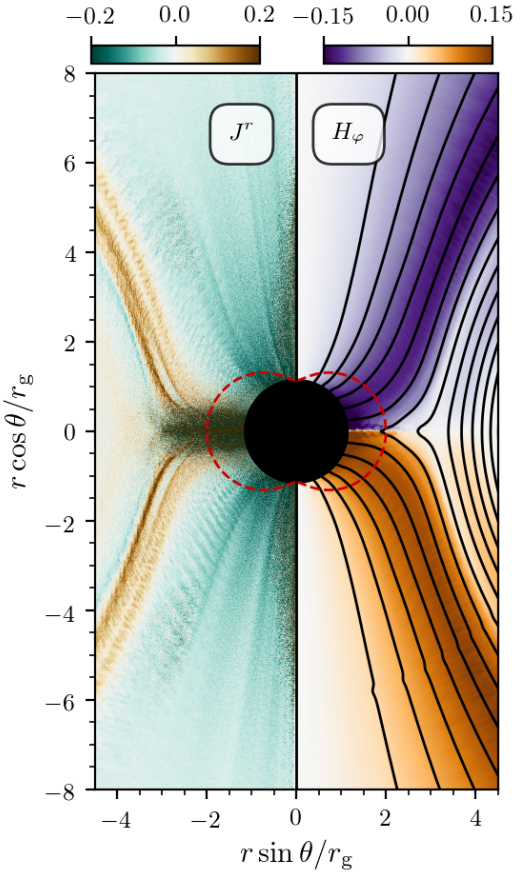


FIGURE 13.2: Left panel: time-averaged normalized electric current density $J^r/(cen_{\mathrm{GJ}})$. Right panel: time-averaged normalized toroidal component H_φ/B_0. Poloidal magnetic field lines are represented by black solid lines. The dashed red line indicates the surface of the ergosphere.

current system. This positive current flows along the separatrix, *i.e.* the last magnetic field line connected to the black hole, which defines the magnetospheric boundary.

The equatorial current sheet is prone to the plasmoid instability, which mediates fast magnetic reconnection. More details on fast reconnection are provided in the Appendix 13.A. Magnetic energy is dissipated and deposited into particles, leading to intense pair creation. Fig. 13.3 shows a snapshot of the photon density above the pair-creation threshold and particle density, both in logarithmic scale. The mechanism described in Chap. 12 is still operating in this new configuration. Bursts of pair creation occur in an intermittent manner near the inner light surface at intermediate latitudes. This fresh plasma mostly follows the magnetic field lines, therefore it mainly flows close to the magnetospheric boundary. Inside the bursts, the plasma density is marginally denser than the Goldreich-Julian density, and the outflowing plasma is highly magnetized. Pair creation is almost quenched near the rotation axis. We checked that in this zone the 4-current is null, although it is spacelike near the horizon at intermediate latitudes. In addition, the acceleration of particles in the X-points of the current sheet triggers pair creation and high-energy photon emission. The plasma density can reach $10^3 n_{\mathrm{GJ}}$ in the current sheet plasmoids.

The magnetosphere displays an interesting dynamical phenomenon, responsible for the replenishment of the magnetic field threading the black hole, which is represented in Fig. 13.4. Starting from an initial state similar to that shown in Fig. 13.2, plasma accumulates near the Y-point of the magnetosphere (defined as the point of the current sheet most distant to the black hole, where the two separatrices diverge from the equator). This plasma is supported

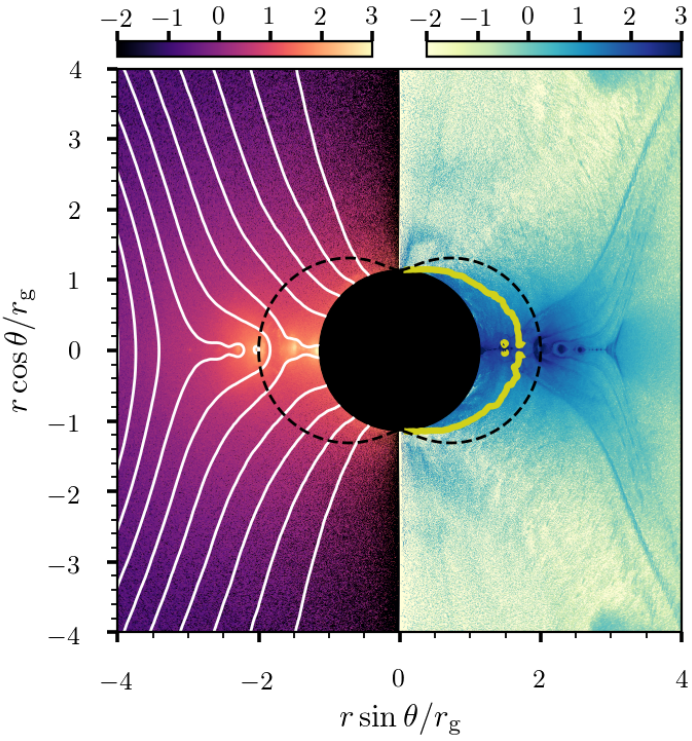


FIGURE 13.3: Snapshot of the simulation with $\tau_0 = 50$ and $V_0/c = 0.05$. Left panel: Logarithm of the density of photons above the threshold, normalized by $n_{\rm GJ}$. Poloidal magnetic field lines are represented by white solid lines. Right panel: Logarithm of the plasma density, normalized by $n_{\rm GJ}$. The dashed black line indicates the surface of the ergosphere, and the yellow solid line on the right panel shows the inner light surface.

by the magnetic pressure inside the magnetosphere. When the magnetosphere can no longer sustain the plasma, which roughly occurs when the particle energy density exceeds the magnetic energy density, a giant plasmoid forms and suddenly plunges into the black hole. This corresponds to the breakdown of the force-free approximation. The weakly magnetized plasma plunges due to the gravitational pull of the black hole, and works against the magnetic tension of field lines. As this giant plasmoid rushes inward, it pulls inward vertical magnetic field lines that were not crossing the event horizon initially. This replenishes the magnetic flux of the black hole. After the black hole swallows the giant plasmoid, the magnetosphere goes back to its initial state, until a new giant plasmoid is formed.

13.2.2 Long-term evolution

The outcome of the simulation is shown by the black and blue curves in Fig. 13.5, which represent the evolution of the magnetic flux Φ through the upper hemisphere of the event horizon with time. The magnetosphere experiences the dynamic cycles described in the previous section for about $300\,r_{\rm g}/c$, but the magnetic flux Φ decays secularly. It settles at a steady value after a time $\simeq 500\,r_{\rm g}/c$. The steady state of the simulation resembles the Wald setup described by Parfrey et al. (2019), as can be seen in Fig. 13.6. The field lines are much more vertical and, more importantly, the Y-point is located very close to the boundary of the ergosphere. In this steady state, there are no more giant plasmoid accretion cycles. The current sheet is still disrupted by the tearing instability, so that small plasmoids fall toward the black hole. The magnetic flux escape by magnetic reconnection is exactly balanced by the supply of magnetic flux caused by the inflowing plasmoids pulling vertical

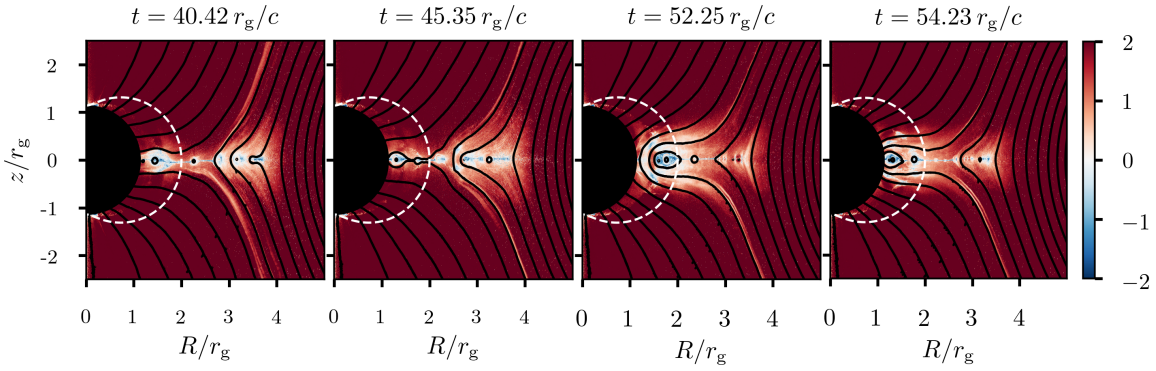


FIGURE 13.4: Snapshots of the logarithm of $\sigma = B^2/\left(8\pi n m_e c^2 \Gamma\right)$ from a simulation with $\tau_0 = 50$. The zones with $\sigma \simeq 1$ are in white. Poloidal magnetic field lines are represented by black solid lines. The dashed white line indicates the surface of the ergosphere.

field lines. Therefore, without external forcing, the only stable configuration for the magnetosphere is Wald-like. This is reminiscent of pulsar magnetospheres, where the Y-point naturally migrates toward an equilibrium position at the light cylinder (Spitkovsky, 2006). This configuration is close to force-free, except in the current sheet.

It should be noted that because the magnetic field strength has dropped significantly, the maximum Lorentz factor particles can achieve is no longer much larger than the pair-creation threshold. Being slightly starved, large gaps can momentarily open up. This is merely an effect of our limited scale separation, and does not affect the general conclusion. We also note that the final value for the magnetic flux does not depend on τ_0, in the range of parameters we have tested. If the opacity of the medium is high enough, the magnetosphere can reach a state close to force-free, irrespective of the plasma supply details.

We also ran simulations with no conducting disk. In these simulations, the equatorial current sheet quickly extends across the whole box. The initial magnetic energy is quickly dissipated, whereas the mechanism previously analyzed cannot take place: there is no available vertical magnetic flux that could compensate for this decay. The simulation is exhausted of particles and dies out (see Sec. 13.4).

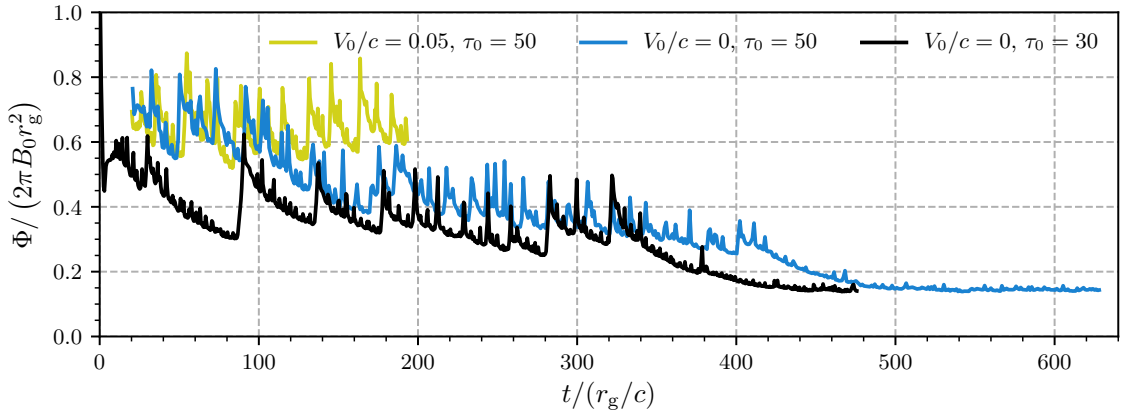


FIGURE 13.5: Evolution of the magnetic flux $\Phi/\left(2\pi B_0 r_g^2\right)$ through the upper hemisphere of the event horizon, for different V_0 and τ_0.

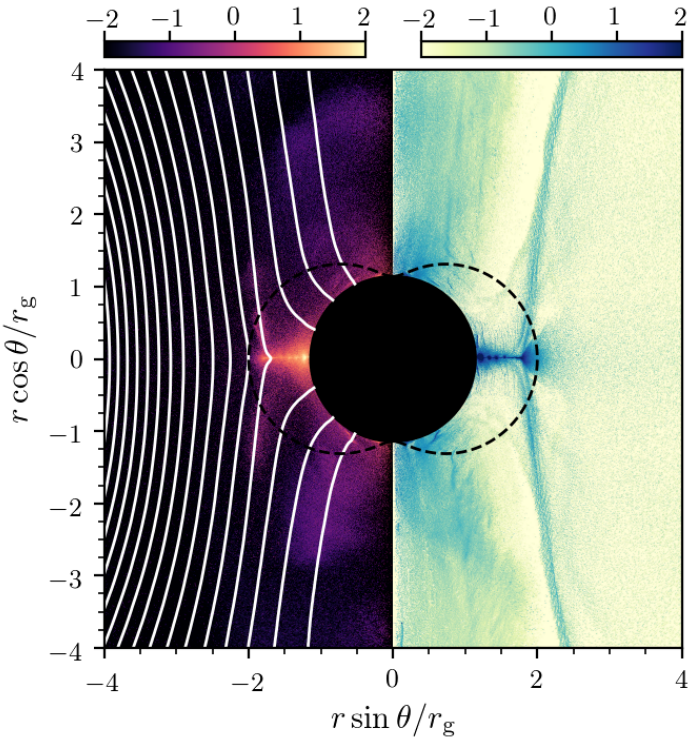


FIGURE 13.6: Steady state of the simulation with $\tau_0 = 50$ and $V_0 = 0$. Left panel: Logarithm of the photon density, normalized by $n_{\rm GJ}$. Poloidal magnetic field lines are represented by white solid lines. Right panel: Logarithm of the plasma density, normalized by $n_{\rm GJ}$. The dashed black line indicates the surface of the ergosphere.

13.3 Magnetic field transport

We are interested in maintaining the dynamic state and impeding magnetic field decay, since this variable state is promising for the prospect of high-energy flares. Therefore, we have added the possibility of supplying magnetic flux to the central black hole, in order to study the response of the magnetosphere either free or forced. To this end, we do not inject magnetic flux in the whole simulation box, but rather advect the frozen-in field lines that are initially crossing the perfectly conducting disk. We add a small toroidal electric field $\boldsymbol{E}_{\rm acc} = -(\boldsymbol{V}_0/c) \times \boldsymbol{B}$ only in the conducting disk. In concrete terms, this amounts to enforcing

$$E_\varphi = \frac{V_0}{c} B^\theta_{\rm in} \tag{13.8}$$

inside the disk (which also impacts how D^θ is prescribed). We run another set of simulations with various values of τ_0, with $V_0/c = 0.05$ this time. This setup can mimic inward magnetic flux transport in accretion flows (Lubow et al., 1994). We use a value of V_0 that is consistent with ideal MHD simulations of accretion disks (Jacquemin-Ide et al., 2021).

By virtue of Faraday's law, applied to a loop of radius $r = r_{\rm in}$ in the equatorial plane (see Eq. (13.10)), the magnetic flux through a surface enclosed by this loop must increase steadily for $V_0 \neq 0$, and remain constant for $V_0 = 0$. In other words, magnetic field lines that have been transported below $r = r_{\rm in}$ at $\theta = \pi/2$ must remain below $r_{\rm in}$ from then on. This is why we place the inner boundary of the conducting disk at sufficient distance $r_{\rm in} = 6\,r_{\rm g}$ from the black hole. We choose not to run simulations with $V_0 \neq 0$ for as long as simulations with $V_0 = 0$, because the magnetic flux within $r = r_{\rm in}$ would accumulate near the ergosphere.

This would not occur in real systems, as there would be a way for the magnetic flux to flow back from the magnetosphere to the disk.

In the presence of magnetic field line transport ($V_0 \neq 0$), the magnetosphere is able to remain in a dynamic state of periodic giant plasmoid accretion events (see the yellow curve in Fig. 13.5). The inflow of magnetic flux compresses the magnetosphere and compensates the secular decay. The evolution of Φ in this dynamic state is represented in the upper panel in Fig. 13.8 for different optical depths, with the four black dots representing successive snapshots relative to Fig. 13.4. As a magnetized giant plasmoid is swallowed by the black hole, the magnetic flux experiences a sharp rise. Between two successive giant plasmoid accretion events, the magnetic flux decays almost exponentially with time due to magnetic reconnection. We observed that the characteristic decay time of Φ barely depends on τ_0. On the other hand, the frequency of these accretion cycles is controlled by the fiducial optical depth τ_0: it increases with increasing τ_0. This is because mass loading at the Y-point is more efficient at high optical depths, which results in more frequent cycles of accretion. These cycles are illustrated in Fig. 13.7, which represents a spacetime diagram of the flux function A_φ in the equatorial plane. They occur with a period of around $15\,r_\mathrm{g}/c$. The slow transport of magnetic field lines from the conducting disk to the black hole is also visible between $3\,r_\mathrm{g}$ and $4\,r_\mathrm{g}$.

The lower panel in Fig. 13.8 shows the time evolution of the Poynting flux through the event horizon for the simulations with magnetic field transport, computed using Eq. (12.5) and (12.6). We also observe sharp rises in the Poynting flux, synchronized with those in Φ. This comes as no surprise, since the output power is expected to scale as Φ^2 if the Blandford-Znajek process is activated (see Eq. (7.20)). In the case of a pure split-monopole magnetosphere, the total Poynting luminosity is still given by Eq. (7.18). We find that the measured luminosity is lower than this estimate. This is because some flux is removed from the event horizon during an initial transient, due to the initial conditions not being an equilibrium state. This also explains why Φ is consistently below $2\pi {r_\mathrm{g}}^2 B_0$.

It is likely that these results would slightly differ in 3D, since nonaxisymmetric modes would also allow interchange of tenuous magnetospheric plasma with dense, unmagnetized plasma at the Y-point. In particular, it is possible that Φ would not experience sharp and periodic peaks. Nonetheless, this mechanism should still hold and allow the black hole to retain a significant magnetic flux and luminosity on a timescale longer than the characteristic reconnection decay time T.

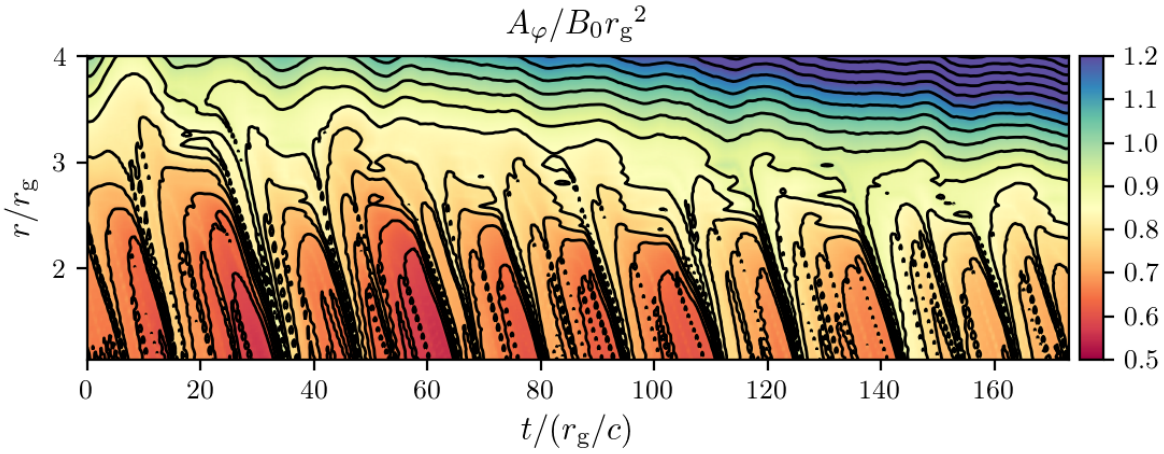


FIGURE 13.7: Spacetime diagram of A_φ in the equatorial plane $\theta = \pi/2$ for the simulation with $\tau_0 = 50$ and $V_0/c = 0.05$. The black solid lines are the contours of A_φ, which represent poloidal magnetic field lines.

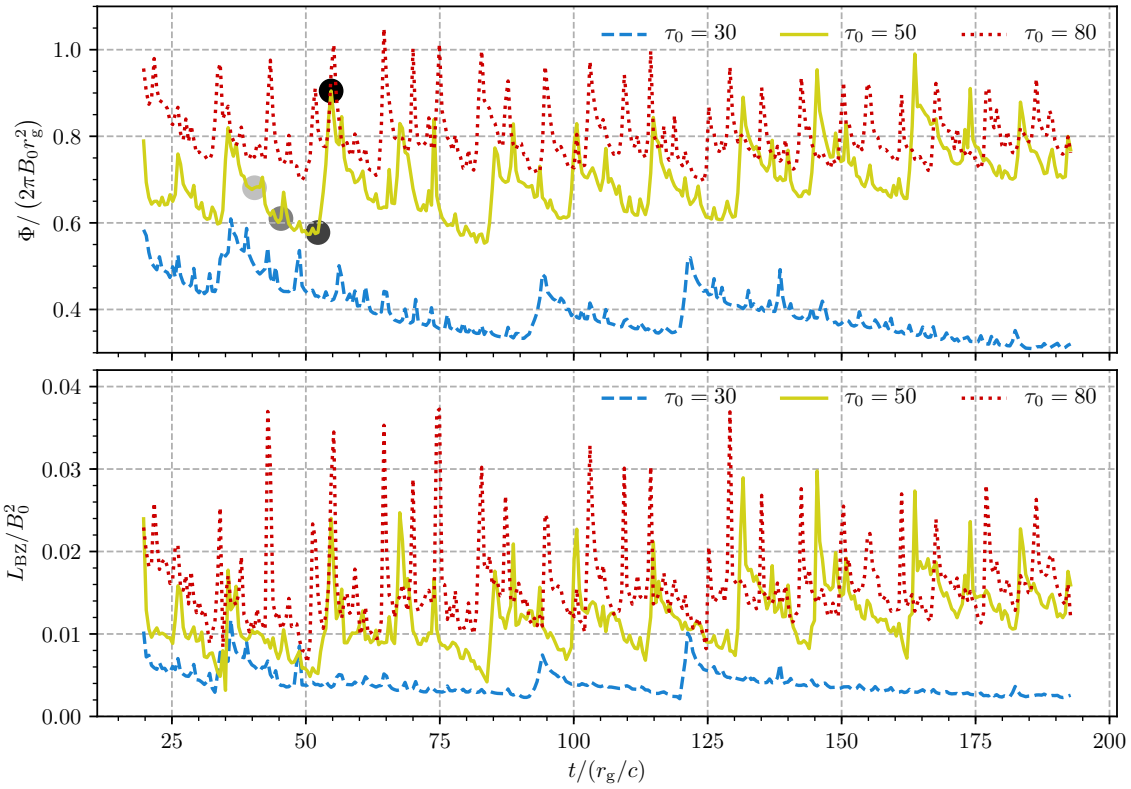


FIGURE 13.8: Upper panel: time evolution of the normalized magnetic flux $\Phi/\left(2\pi B_0 r_{\rm g}^2\right)$ through the upper hemisphere of the event horizon, for the three simulations at $V_0/c = 0.05$. The blue disks, relative to the $\tau_0 = 50$ simulation, indicate the times of the snapshots relative to Fig. 13.4. Lower panel: time evolution of the normalized Poynting flux $L_{\rm BZ}/B_0^2$, integrated over the event horizon, for these same three simulations.

13.4 Toy model for magnetic flux decay

The decay of the magnetic flux Φ through the event horizon, in the absence of any source term due to inflowing plasmoids, is a consequence of dissipation of magnetic energy by magnetic reconnection (see the Appendix 13.A). Let us provide a toy model to account for the order of magnitude of the characteristic decay time T, assuming axisymmetry. Following Eq. (5.40), the magnetic flux Φ through the upper hemisphere of the event horizon can be expressed as

$$\Phi = \iint B^r \sqrt{h}\, {\rm d}\theta\, {\rm d}\varphi = 2\pi A_\varphi(r_{\rm h}, \pi/2). \tag{13.9}$$

Faraday's law allows us to express the time derivative of Φ as the circulation of $\boldsymbol{E}$ along a loop of radius $r_{\rm h}$ in the equatorial plane:

$$\frac{{\rm d}\Phi}{{\rm d}t} = 2\pi \partial_t A_\varphi(r_{\rm h}, \pi/2) = -2\pi c E_\varphi(r_{\rm h}, \pi/2). \tag{13.10}$$

In a purely axisymmetric and stationary magnetosphere, we have $E_\varphi = 0$ everywhere. Yet, this component arises in and near the current sheet because there is an inflow of plasma at velocity $\boldsymbol{V}_{\rm in} = V_{\rm in}\partial_\theta$ toward the reconnection region. Just above and below the current sheet, the electric field reads $\boldsymbol{E} = -\left(\boldsymbol{V}_{\rm in}/c\right) \times \boldsymbol{B}$. At such high magnetizations, the outflow

velocity is very close to c. Thence, we define a global dimensionless reconnection rate R as

$$\mathcal{R} = \frac{|V_{\rm in}|}{c} = \frac{r_{\rm g} E_\varphi}{\sqrt{h}|B^r|}, \tag{13.11}$$

where h is evaluated at $r = r_{\rm h}$ and $\theta = \pi/2$. Here, $V_{\rm in}$ is not measured by the FIDO but with respect to the grid. We also assume that the configuration of field lines at the event horizon is close to a split monopole, so that $B^r(r = r_{\rm h})$ does not depend much on θ in a given hemisphere. The magnetic flux can be written as $\Phi = B^r(r_{\rm h})\mathcal{S}$, where $\mathcal{S}$ is half the area of the event horizon, which we have already computed (see Eq. (5.92)). Ultimately, the evolution of the magnetic flux is governed by

$$\frac{{\rm d}\Phi}{{\rm d}t} = -\frac{2\pi\sqrt{h}}{\mathcal{S}}\frac{c}{r_{\rm g}}\mathcal{R}\,\Phi. \tag{13.12}$$

If the global reconnection rate is time-independent, the magnetic flux decreases exponentially with a characteristic decay time

$$T = \frac{\mathcal{S}}{2\pi\sqrt{h}}\frac{1}{\mathcal{R}}\frac{r_{\rm g}}{c} \simeq \frac{1}{\mathcal{R}}\frac{r_{\rm g}}{c} \tag{13.13}$$

From Fig. 13.8 and 13.5, we measure the slope of the exponential decay, and obtain a reconnection rate $R = 0.02 \pm 0.002$, corresponding to a decay time $T \simeq 50\, r_{\rm g}/c$. We also measured the local reconnection rate using Eq. (13.11). The time-averaged result is shown in Fig. 13.9 for the simulation at $\tau_0 = 30$ and $V_0 = 0$. The reconnection plasma inflow is characterized by $\mathcal{R} \neq 0$, which is true above and below the current sheet. The out-of-plane electric field is generally especially intense at the X points, where the electromagnetic field is electrically dominated. In addition to that, the out-of plane is nonzero in a thick sheath around the reconnection layer, which is visible in Fig. 13.9. This has been observed in local simulations of relativistic magnetic reconnection (Cerutti et al., 2013). We find that typical values of $\mathcal{R}$ range from 0.02 to 0.04, which is consistent with the flux decay time. This model naturally explains why the characteristic decay time is the same in all simulations.

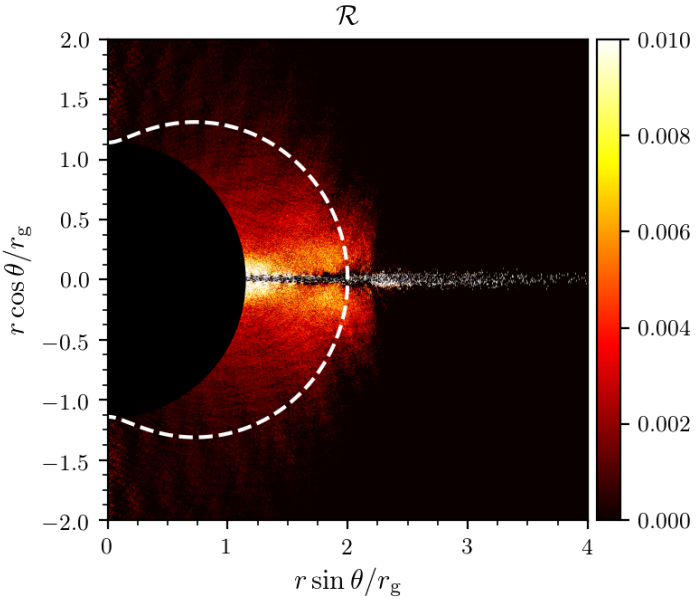


FIGURE 13.9: Time-averaged local reconnection rate $\mathcal{R}$ for $\tau_0 = 30$ and $V_0 = 0$.

The local reconnection rate in collisionless relativistic reconnection in pair plasmas has been determined by several numerical studies (Kagan et al., 2013; Sironi & Spitkovsky, 2014), with typical values between 10^{-2} and 10^{-1}. To compare with the decay time T (measured by an observer at infinity), one needs to take into account gravitational dilation. Assuming the current sheet is roughly comoving (radially inward) with the Kerr-Schild FIDO at $r = r_{\rm h}$ and $\theta = \pi/2$, by definition of the lapse function α, the local reconnection rate can be estimated as $\mathcal{R}/\alpha \simeq 1.66\,\mathcal{R}$. Our value of $\mathcal{R}$ is consistent with measurements of the local collisionless reconnection rate, although slightly low.

13.A Appendix: Relativistic magnetic reconnection

13.A.1 Introduction

Magnetic reconnection is an ubiquitous phenomenon in astrophysical plasmas (see Zweibel & Yamada 2009 and Yamada et al. 2010 for reviews). This phenomenon can be defined as a change in the topology of the magnetic field lines. Because most astrophysical plasmas are highly conductive, magnetic field lines are usually frozen in the plasma and cannot change their topology: this is the regime of ideal MHD. Only if nonideal effects (such as resistivity) are significant can there be magnetic reconnection. In most astrophysical plasmas, the resistivity is so small that other nonideal effects first come into play: electron inertia, Hall effects, finite-Larmor radius effects... The frozen-flux condition breaks down in the reconnection zone. Magnetic reconnection requires multiple spatial scales: a small diffusive scale at which ideal MHD is not valid, and a large region providing plasma. Solar flares or magnetospheric storms are probably powered by magnetic reconnection.

If the plasma magnetization σ, defined as the ratio of the magnetic energy density and the total particle enthalpy density, is much larger than 1, magnetic reconnection is said to occur in the relativistic regime. Most of the magnetic energy is transferred to the plasma, so that the plasma outflow powered by reconnection is relativistic. In this regime, the Alfvén speed of the plasma is close to the speed of light (Lyutikov & Uzdensky, 2003):

$$v_{\rm A} = \sqrt{\frac{\sigma}{1+\sigma}}c \simeq c. \tag{13.14}$$

Relativistic reconnection is thought to operate in pulsar winds (Coroniti, 1990), accreting black-hole coronæ (Goodman & Uzdensky, 2008) or gamma-ray bursts (Drenkhahn & Spruit, 2002) for instance. A review of relativistic magnetic reconnection can be found in Kagan et al. (2015). It is now firmly established that magnetic reconnection is efficient at dissipating magnetic free energy and converting it into thermal or kinetic energy in a fast way. It leads to strong particle acceleration and can produce very hard spectra, in particular in magnetically dominated systems (Sironi & Spitkovsky, 2014; Werner et al., 2016).

13.A.2 Steady-state reconnection

Let us describe a 2D configuration for simplicity. Magnetic reconnection involves a change of polarity in the magnetic field, with a plasma current sheet inbetween. This situation is represented in Fig. 13.10, the reconnecting fields being anti-parallel and of equal strength B_0, along $\boldsymbol{e}_x$. The situation is invariant by translation along the z axis.

The reconnecting field is allowed to diffuse through the current sheet, and reconnects at an X-point. Plasma flows toward the elongated current sheet at low speed and is ejected sideways at the Alfvén speed, as field lines snap and reconnect. The outflow is a consequence of the magnetic tension of the highly bent reconnected field lines. Outside of the

current sheet, the nonideal effects are negligible, so that the electric field $\boldsymbol{E}$ is related to the magnetic field $\boldsymbol{B}$ by

$$\boldsymbol{E} + \frac{1}{c}\boldsymbol{V}_{\rm in} \times \boldsymbol{B} = \boldsymbol{0}, \tag{13.15}$$

with $\boldsymbol{V}_{\rm in}$ the inflow bulk velocity. The electric field is out-of-plane (along $\boldsymbol{e}_z$). In a steady-state configuration, $\boldsymbol{E}$ must be uniform, so that it is also nonzero in the current sheet (where $\boldsymbol{B}$ can vanish), where it is related to nonideal effects. Ultimately, this electric field originates from magnetic flux variations and Faraday's law, as magnetic field lines are pushed toward the current sheet. Reconnection occurs when this electric field can drive out-of-plane currents through nonideal effects, which modify the magnetic topology by inducing magnetic fields in the $\boldsymbol{e}_y$ direction. Whether the electric field first grows in the current sheet or in the ideal region depends on whether reconnection is externally driven or spontaneous. The reconnection rate $\mathcal{R}$ is defined as the ratio of the inflow to the outflow velocity. In the relativistic limit, it is defined as $\mathcal{R} = V_{\rm in}/c$. From Eq. (13.15), we see that it can be recast into $\mathcal{R} = E_z/B_x$.

The canonical reconnection configuration is the Sweet-Parker model (Parker, 1957; Sweet, 1958), in the MHD framework. It was described in the relativistic regime by Lyubarsky (2005). Let us denote δ the width of the current sheet and L its length. The width of the current sheet is determined by nonideal processes and by the reconnection rate, which depends on microphysics. From mass conservation, the reconnection rate is directly given by the aspect ratio of the layer: $V_{\rm in}/c = \delta/L$. In the Sweet-Parker model, the reconnection rate is given by

$$\mathcal{R} = \frac{\delta}{L} = \frac{1}{\sqrt{S}}; \quad S = \frac{4\pi c L}{\eta}, \tag{13.16}$$

with S the Lundquist number and η the resistivity of the plasma. If L is taken to be the typical size of the system, S is usually tremendously high ($S \simeq 10^{20}$ is typical), so that reconnection occurs at a very low rate. This is at odds with observations of solar flares, for instance. Yet, evidence for fast reconnection is conspicuous. If the Sweet-Parker model is right, as laboratory measurements (Ji et al., 1999) and numerical simulations of collisional plasmas (Uzdensky & Kulsrud, 2000) seem to suggest, the length of the reconnecting current sheet must be drastically smaller than the size of the system.

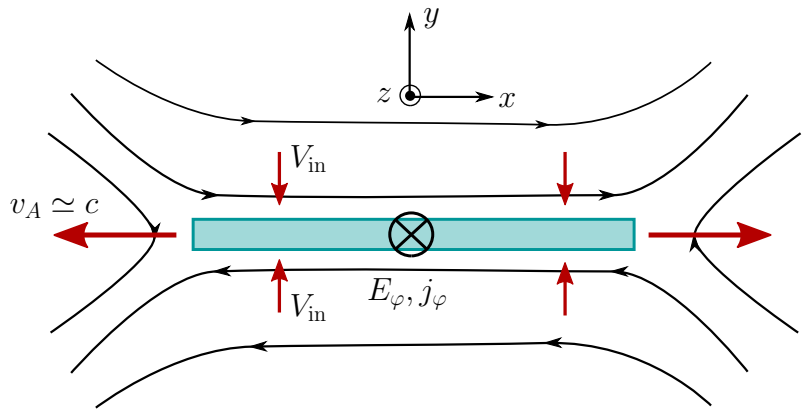


FIGURE 13.10: Sketch of the Sweet-Parker model. Plasma inflows toward the reconnecting layer, of length L and thickness δ, with velocity $V_{\rm in}$. An electric field develops perpendicularly to the plane of the figure. Plasma is ejected sideways with the Alfvén speed.

13.A.3 Fast reconnection

A review of the recent progress on magnetic reconnection can be found in Schekochihin (2020). Let us consider two adjacent regions of opposite magnetic polarities, with a typical spatial scale Δ for the variation of the magnetic field. A current sheet such as the one discussed in the previous section can be the result of external forcing, with a plasma flow pushing the two magnetic polarities together, but it can also arise spontaneously. In the absence of forcing, this magnetic configuration is liable to the tearing instability (Schekochihin, 2020). In the nonlinear stage of this instability, magnetic islands of size $\simeq \Delta$ separated by X-points are formed. These X-points are then found to collapse (Loureiro et al., 2005), leading to the formation of a current sheet with sideways Alfvénic outflows.

Let us assume that a current sheet is readily formed, irrespective of the outer environment. Sweet-Parker reconnection could *a priori* occur at a rate given by Eq. (13.16), but it was realized that the aspect ratio of the sheet δ/L could not be arbitrarily small. In the framework of resistive MHD, analytical calculations (Loureiro et al., 2007) and numerical simulations (Samtaney et al., 2009) showed that a current sheet with small aspect ratio broke up into a chain of magnetic islands, or "plasmoids". They also found that the critical aspect ratio was $\delta/L \simeq 10^{-2}$, corresponding to $S \simeq 10^4$. If $S < 10^4$, the Sweet-Parker current sheet was found to be stable. Numerical simulations of collisionless plasmas also indicate a similar trend, with the critical aspect ration being $\delta/L \simeq 0.1$ (Sironi et al., 2016). Interestingly, this plasmoid instability has a growth rate much larger than the characteristic outflow time L/c, so that secondary current sheets are formed between the plasmoids.

From these observations, Uzdensky et al. (2010) proposed the following model of fast magnetic reconnection. A current sheet with an aspect ratio beyond the critical aspect ratio breaks up into plasmoids, a process which happens sufficiently fast for the nonlinear stage of the plasmoid instability to develop before the plasmoids are ejected. Secondary current sheets are formed between these plasmoids. Plasmoids are ejected sideways, stretching the current sheets and making them plasmoid-unstable themselves. A hierarchy of plasmoids can be found in the reconnecting region. The process repeats until the smallest secondary current sheets are sufficiently short that they are no longer plasmoid-unstable. At this stage, Sweet-Parker reconnection can proceed. The reconnection rate becomes independent of the global Lundquist number, and only depends on the critical aspect ratio. Hence, reconnection occur in a relatively fast regime, with a global effective rate close to 0.01 for resistive MHD and of the order 0.1 for collisionless plasmas. These values are measured in various regimes and configurations, and can be regarded as "universal".

Chapter 14
Synthetic observables

As discussed previously, high-energy photons below the pair creation threshold are discarded from the simulation, as they do not participate in the simulation. Yet, in order to reconstruct the radiative signature of the magnetosphere, it is crucial to save the information that they carry. Given their initial positions, directions and emission times, my goal is to reconstruct light curves for different viewing angles (with respect to the spin axis). This task has already been performed in flat spacetime, in the context of pulsar magnetospheres (Cerutti et al., 2016). Here, this approach must be generalized to photons propagating in a curved spacetime.

Beyond a given radius $r_{\rm out}$, photons can be considered to propagate in straight lines. However, relativistic effects cannot be neglected, as the photons are emitted close to the black hole. In order to compute their final directions, times of flight, and impact parameters, it is compulsory to integrate the null geodesics of the photons from their emission points up to $r = r_{\rm out}$. Keeping these photons in the simulation box and integrating their equations of motion with the PIC algorithm, even with a looser constraint on the time step, would be too demanding computationally. As it happens, there is no need to solve the entire geodesic, since the only relevant information is the initial and final coordinates (t, r, θ, φ) of the photons. Instead, I have used the public ray-tracing code `geokerr` (Dexter & Agol, 2009), which is optimized to numerically integrate null geodesics in the Kerr metric.

This method has allowed me to build both light curves and images. The former are relevant in the context of VHE emission. Moreover, photons emitted by synchrotron process can be a good proxy for the millimeter emission from M87*. Our goal is to model the optically thin millimeter emission, in order to evaluate the contribution of magnetospheric processes to the image observed by the EHT. Indeed, the close environment of black holes cannot be captured by GRMHD simulations. This is exemplified by the fact that in the post-processing radiative transfer models used to interpret the image, the funnel is cut out (Event Horizon Telescope Collaboration et al., 2019b, 2021b). It is not clear what is the main emission sites of the photons that make up the image.

In Sec. 14.1, I briefly describe the theory of null geodesics in Kerr spacetime and how I coupled `GRZeltron` with `geokerr`. The resulting light curves and images are presented in Sec. 14.2 and 14.3 respectively. The section on the light curves is adapted from Crinquand et al. (2021), whereas the results on the images are only preliminary and have not been published yet.

14.1 Null geodesics in Kerr spacetime

In this section, we employ BL coordinates.

14.1.1 Equations of motion

Since the wavelength of the light of interest is much smaller than the radius of the black hole, we can work in the geometric optics limit, so that photons follow null geodesics. Let

us consider a photon with 4-momentum p_μ. Geodesic motion in Kerr spacetime possesses the pleasant property of being completely integrable, since three conserved quantities can be attributed to each geodesic. These constants are:

- the energy $E = -p_t = -cp_0$;
- the angular momentum $L = p_\varphi$;
- the Carter constant Q (Carter, 1968), defined as

$$Q = {p_\theta}^2 - \cos^2\theta \left(a^2 \frac{E^2 r_\mathrm{g}^2}{c^2} - \frac{L^2}{\sin^2\theta} \right). \tag{14.1}$$

The first two constants are related to the Killing vectors of the Kerr spacetime (see Sec. 4.2.1). These equations are supplemented with the massless constraint $p_\mu p^\mu = 0$, which makes geodesic motion fully integrable by providing a fourth equation. It is convenient to define the dimensionless parameters $\ell = cL/Er_\mathrm{g}$ and $q^2 = Qc^2/E^2 r_\mathrm{g}^2$. The Carter constant is related to the orthoradial momentum p_θ, but is not always positive. If a photon reaches spatial infinity, E is its energy as measured by an observer at infinity, which is redshifted and differs from the emitted photon energy.

> The existence of this third Carter constant is a stroke of luck, which is not related to any obvious symmetry of spacetime.

From the constants of motion and the constraint $p_\mu p^\mu = 0$, the contravariant components of the photon momentum can be reconstructed (Bardeen et al., 1972):

$$\frac{\Sigma}{r_\mathrm{g}^2}\frac{cp^r}{E} = \pm\sqrt{\mathcal{R}(r)}, \tag{14.2}$$

$$\frac{\Sigma}{r_\mathrm{g}}\frac{cp^\theta}{E} = \pm\sqrt{\Theta(\theta)}, \tag{14.3}$$

$$\frac{\Sigma}{r_\mathrm{g}}\frac{cp^\varphi}{E} = -\left(a - \frac{\ell}{\sin^2\theta}\right) + a\frac{r^2 + a^2 r_\mathrm{g}^2 - a r_\mathrm{g}^2 \ell}{\Delta}, \tag{14.4}$$

$$\frac{\Sigma}{r_\mathrm{g}^2}\frac{cp^0}{E} = -a\left(a\sin^2\theta - \ell\right) + \frac{r^2 + a^2 r_\mathrm{g}^2}{\Delta r_\mathrm{g}^2}\left(r^2 + a^2 r_\mathrm{g}^2 - a r_\mathrm{g}^2 \ell\right), \tag{14.5}$$

where the radial potential $\mathcal{R}$ and angular potential Θ are given by

$$\mathcal{R}(r) = \left(r^2/r_\mathrm{g}^2 + a^2 - a\ell\right)^2 - \frac{\Delta}{r_\mathrm{g}^2}\left(q^2 + (a-\ell)^2\right), \tag{14.6}$$

$$\Theta(\theta) = q^2 + a^2\cos^2\theta - \ell^2 \operatorname{cotan}^2\theta. \tag{14.7}$$

This can be easily generalized to timelike geodesics, with $p_\mu p^\mu = -m^2c^2$.

From these equations, it is readily seen that geodesic motion does not depend on the energy (or the wavelength) of the photon: gravity is "achromatic". Note that Eq. (14.2) and (14.3) are defined up to a choice of sign. This means that for a given set of constants of motion (ℓ, q^2), two geodesics are associated with these parameters (determined by the choice of one of these two signs), whereas the other choice of sign determines the direction of travel along that geodesic. These contravariant components p^μ are actually the derivatives $p^\mu = \mathrm{d}x^\mu/\mathrm{d}\lambda$, where λ is an affine parameter along the geodesic. Therefore, the zeros of the potentials $\mathcal{R}$

and Θ correspond to turning points in the trajectories, with a subsequent sign flip in (14.2) and (14.3).

14.1.2 Impact parameters

Let us assume that a geodesic with parameters (ℓ, q^2) reaches spatial infinity. In order to reconstruct observables, it is very useful to look at its behavior at spatial infinity. Let us assume that the photon is captured by a remote detector located at a point $(r_{\rm f}, \theta_{\rm f}, \varphi_{\rm f} = 0)$, with $r_{\rm f} \gg r_{\rm g}$. In this limit, Eq. (14.3), (14.4) and (14.5) become:

$$\frac{cp^\theta}{E} \simeq \frac{r_{\rm g}}{r_{\rm f}^2}\sqrt{q^2 + a^2\cos^2\theta_{\rm f} - \ell^2\,\mathrm{cotan}^2\,\theta_{\rm f}}; \quad \frac{cp^\varphi}{E} \simeq \frac{\ell r_{\rm g}}{r_{\rm f}^2\sin^2\theta_{\rm f}}; \quad \frac{cp^0}{E} \simeq 1. \tag{14.8}$$

From this, the rates of change of φ and θ at the observers's location can be obtained:

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = c\frac{p^\theta}{p^0} \simeq \pm c\frac{r_{\rm g}}{r_{\rm f}^2}\sqrt{q^2 + a^2\cos^2\theta_{\rm f} - \ell^2\,\mathrm{cotan}^2\,\theta_{\rm f}}; \quad \frac{\mathrm{d}\varphi}{\mathrm{d}t} = c\frac{p^\varphi}{p^0} \simeq c\frac{r_{\rm g}}{r_{\rm f}}\frac{\ell}{r_{\rm f}\sin^2\theta_{\rm f}}. \tag{14.9}$$

A quantity of interest is the final 3-velocity $v^{\hat{i}}$ of the photon in an orthonormal basis, since $\boldsymbol{u} = \boldsymbol{v}/c$ is the unit vector tangent to the photon trajectory. For $r_{\rm f} \gg r_{\rm g}$, we get from Eq. (14.9):

$$v^{\hat\theta} = r_{\rm f}\frac{\mathrm{d}\theta}{\mathrm{d}t} = \pm c\frac{r_{\rm g}}{r_{\rm f}}\sqrt{q^2 + a^2\cos^2\theta_{\rm f} - \ell^2\,\mathrm{cotan}^2\,\theta_{\rm f}}, \tag{14.10}$$

$$v^{\hat\varphi} = r_{\rm f}\sin\theta_{\rm f}\frac{\mathrm{d}\varphi}{\mathrm{d}t} = c\frac{r_{\rm g}}{r_{\rm f}}\frac{\ell}{\sin\theta_{\rm f}}, \tag{14.11}$$

$$v^{\hat r} = \pm\sqrt{c^2 - v^{\hat\theta 2} - v^{\hat\varphi 2}}. \tag{14.12}$$

Obviously, if the photon makes it to infinity, its velocity is bound to become purely radial, so that $v^{\hat r} \to c$ and $v^{\hat\theta}, v^{\hat\varphi} \to 0$ as $r_{\rm f} \to +\infty$.

Now let us assume that a distant observer captures photons reaching a spherical screen $r_{\rm f} = r_{\rm out}$, and aims to determine the photon's outgoing direction, and to produce an image of all the photons propagating in a given direction α. The observer is equipped with a screen, represented in Fig. 14.1(a), the center of which lies on the black hole. The screen is perpendicular to the line of sight of the distant observer to the black hole, which has an inclination α with respect to the spin axis (see Fig. 14.1(b)). It is tangent to the spherical screen centered on the black hole. The impact parameters of a given photon on this screen are denoted as (X, Y). The impact parameter Y is the apparent displacement of the photon in the direction parallel to the spin axis of the black hole (the projected arrow in Fig. 14.1(a)), whereas X is the apparent displacement in the direction perpendicular to the projected axis.

The impact parameters of a given photon parametrized by (ℓ, q^2) can be obtained geometrically (Cunningham & Bardeen, 1973; Frolov & Zelnikov, 2011). An algebraic proof is also given by Gralla et al. (2018). This is illustrated in the simple case of a photon with $\ell = 0$ in Fig. 14.1(b). If $r_{\rm f} \gg r_{\rm g}$, the angle at which the photon reaches the external screen $\theta_{\rm f}$ is very close to α, which is the angle that its outgoing direction $\boldsymbol{u}$ makes with the spin axis. The apparent displacement Y can be related to the small orthoradial velocity of the photon. Defining the angle $\chi = \alpha - \theta_{\rm f}$, we have $\sin\chi = Y/r_{\rm f} = u^{\hat\theta} = v^{\hat\theta}/c$ (see Fig. 14.1(b)), so that the impact parameter reads

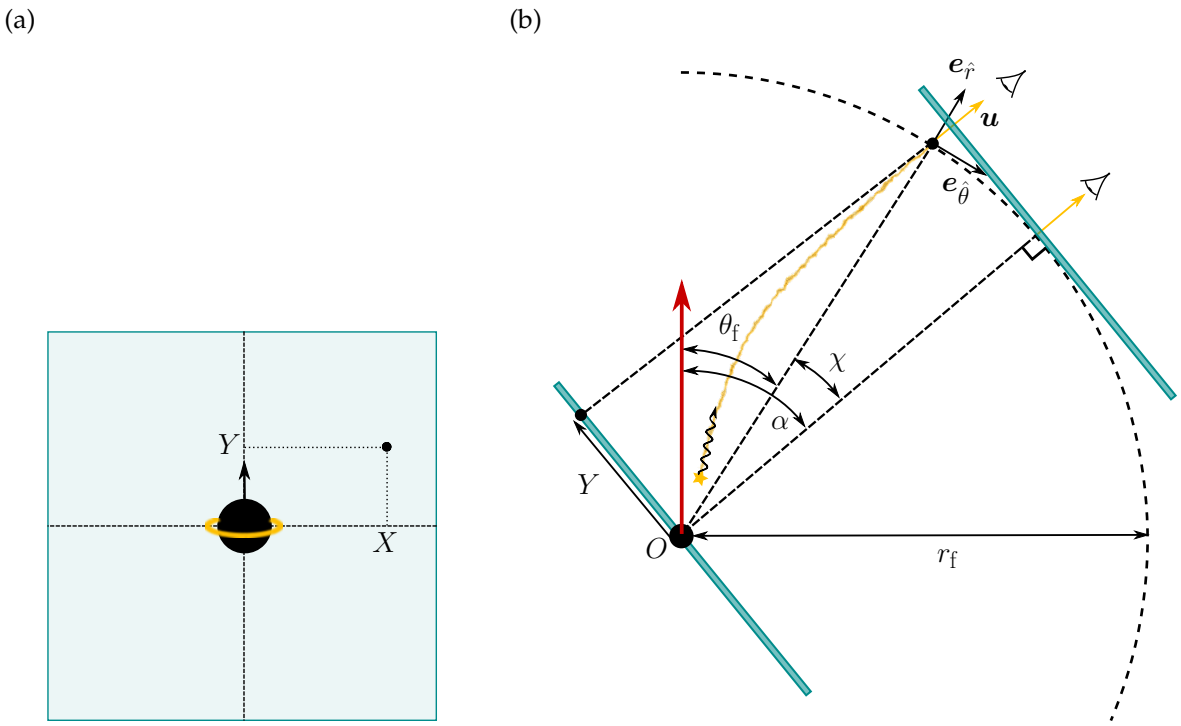


FIGURE 14.1: (a) Observer's screen, centered on the black hole (black disk). The disk is seen face-on, and is normal to the line of sight of the observer, with inclination α. The arrow represents a projection of the black-hole spin, which defines the Y direction. The yellow ring represents the equatorial plane of the black hole, in order to illustrate its inclination. Figure inspired by Cunningham & Bardeen (1973). (b) Situation seen edge-on. The blue rectangle is the observer's screen as seen from the edge. A photon is emitted in the vicinity of the black hole (at the yellow star) and reaches the distant spherical screen $r = r_f$ at $\theta = \theta_f$. This photon escapes toward the observer, its direction $\boldsymbol{u}$ making an angle α with the spin axis (red arrow). The magnitude of the angle χ has been exaggerated for illustrative purposes.

$$\frac{Y}{r_g} \simeq \pm\sqrt{q^2 + a^2\cos^2\alpha - \ell^2\,\mathrm{cotan}^2\,\alpha} = \pm\sqrt{\Theta(\alpha)}. \tag{14.13}$$

Recall that there is freedom in the chosen sign for the orthoradial direction. The same line of reasoning can be carried out for the X parameter, which is related to the azimuthal velocity by $X/r_f = v^{\hat{\varphi}}/c$, so that we have

$$\frac{X}{r_g} = -\frac{\ell}{\sin\alpha}. \tag{14.14}$$

The minus sign comes from the reversed view of the observer, just like East and West are reversed on a sky chart. All in all, for a given photon detected at (r_f, θ_f) with known parameters (ℓ, q^2), as well as two signs defining whether θ and r initially increase or decrease, it is possible to retrieve the impact parameters at a given viewing angle α.

14.1.3 Ray tracing codes

Because accretion flows in low-luminosity AGN are optically thin, the observed light can originate from the innermost regions, where relativistic effects are important. General relativistic ray tracing must be carried out in order to trace the light from its emission region to the observer and to compute the observed intensity. There currently exists a variety of ray

tracing codes in Kerr spacetime, or in any input metric (Vincent et al., 2011; Kuchelmeister et al., 2012; Chan et al., 2013). Most of these perform backward ray tracing. Photons are sent from the observer's screen and integrated backward in time, to the emitting region. This is usually very efficient, as most photons emitted by the object do not end up hitting the observer's screen. Unfortunately, here we track photons *in situ*, so that we must resort to forward ray tracing. I chose to use the public ray tracing code `geokerr`, designed by Dexter & Agol (2009). This code performs geodesic integration in Kerr spacetime only, which is enough for my purpose. This restriction allowed its developers to use a semi-analytic approach, by reducing the equations of motion to tabulated integrals. Thus, this ray tracing code shows excellent performance, both in speed and accuracy.

The procedure goes as follows. A given photon trajectory is defined by the values of ℓ, q^2, the initial radius r_0, the final radius r_f, the initial angle θ_0, and the initial signs of $\mathrm{d}r/\,\mathrm{d}t$ and $\mathrm{d}\theta/\,\mathrm{d}t$. It is also possible to prescribe the final angle instead of the final radius. Integrating Eq. (14.2) and (14.3) eliminates the affine parameter λ, so that the geodesic equation reduces to quadratures (Cunningham & Bardeen, 1973; Gralla et al., 2018):

$$\int_{r_0}^{r_\mathrm{f}} \frac{\mathrm{d}r/r_\mathrm{g}}{\pm\sqrt{\mathcal{R}(r)}} = \int_{\theta_0}^{\theta_\mathrm{f}} \frac{\mathrm{d}\theta}{\pm\sqrt{\Theta(\theta)}}. \tag{14.15}$$

This integral should be understood as a line integral along the trajectory, with sign flips when the geodesic arrives at a turning point in r or θ. This is the equation that the geodesic connecting the point (r_0, θ_0) to $(r_\mathrm{f}, \theta_\mathrm{f})$ must verify, allowing to determine θ_f. Knowing both r_f and θ_f, the variations of φ and t can then be obtained by integrating Eq. (14.4) and (14.5). Since all equations of motion are handled in an integrated form, the need to "spatially resolve" the geodesic is relaxed. Indeed, all we need to compute the times of flight and the impact parameters is the final angle θ_f. Consequently, this procedure is very efficient in our case: we do not need to compute a large number of points on the geodesic. This is unlike radiative transfer codes, which integrate the emissivity of the medium surrounding the black hole and must resolve the entire geodesic.

14.1.4 Coupling with `geokerr`

Let us describe how to use `geokerr` for our purposes, first focusing on gamma rays. At every time step, if a high-energy photon produced by inverse Compton scattering is measured to be under the pair creation threshold, there is a probability f that its information is dumped to an external file. This file is later to be processed by `geokerr`. Then, the photon is discarded. That way, we save the information of a fraction $f = 10\%$ of photons below the threshold. To save computation time, we determine in advance whether a photon travels on a geodesic onto spatial infinity or onto the black hole. Our goal is to only dump photons that will reach the observer at infinity. This can be done by determining the roots of the radial potential $\mathcal{R}$, as explained in the block below.

Method

$\mathcal{R}$ is a quartic polynomial with no third power of r. One can check that $\mathcal{R}(r_\mathrm{h}) > 0$, $\mathcal{R}(0) < 0$, and that $\mathcal{R}$ is positive at $\pm\infty$. As a result, it has at least two real roots and no more than four real roots. Because the sum of all four roots must be zero and $\mathcal{R}(r_\mathrm{h})$ is positive, there has to be two real roots smaller than r_h, which have no physical significance. If the two remaining roots are complex, the photon trajectory extends from r_h to infinity. If they are real, denoted as r_1 and $r_2 \geq r_1$, they lie in $[r_\mathrm{h}, +\infty[$. The different cases depend on the existence of the real roots r_1 and r_2, the photon's initial position r_0 and its initial velocity v^r.

1. There are no real roots in $[r_\mathrm{h}, +\infty[$. The photon reaches infinity only if $v^r > 0$.
2. There are two real roots $r_1 \leq r_2$ in $[r_\mathrm{h}, +\infty[$.
   (a) If $r_0 < r_1$, the photon falls to the black hole, even if $v^r > 0$.
   (b) If $r_0 > r_2$, the photon escapes to infinity, even if $v^r < 0$.

In any case, the radial motion has no more than one turning point (Dexter & Agol, 2009; Frolov & Zelnikov, 2011).

The code evolves the position x^i and components of the momentum p_i of a given photon according to Eq. (10.72) and (10.73). From Eq. (10.71), the energy-at-infinity E can be deduced from x^i and p_i as

$$E = \alpha(x^i)\,\mathcal{E} - c\beta^i(x^i)\,p_i, \tag{14.16}$$

where $\mathcal{E} = c\sqrt{h^{jk}p_j p_k}$ is the FIDO-measured energy. This allows us to compute the dimensionless numbers $\ell = cp_\varphi/r_\mathrm{g}E$ and q^2 (Eq. (14.1)). Each photon dumped in the `geokerr` input file is characterized by the following information:

- the parameters ℓ and q^2,
- the initial position r_0 and θ_0,
- the terminal radius r_f, chosen at a spherical screen $r_\mathrm{f} = r_\mathrm{out} = 200\,r_\mathrm{g} \gg r_\mathrm{g}$,
- the initial signs of $\mathrm{d}r/\,\mathrm{d}t$ and $\mathrm{d}\theta/\,\mathrm{d}t$,
- the weight of the particle and its energy-at-infinity E,
- the initial time of emission t_0 and azimuthal angle φ_0.

The last two items are not needed by `geokerr`, but are used to reconstruct the light curves and images. This code is written in BL coordinates, unlike `GRZeltron` which works in KS coordinates. As a result, when dumping t_0 and φ_0, the values from `GRZeltron` should be converted using Eq. (4.25a) and (4.25b). This procedure provides us a list of photons with all the necessary information to reconstruct their full trajectories. All diagnostics can be performed in post-processing. Then, `geokerr` is run on this input file. The output file provides, for each photon, the terminal angle θ_f at which the photon reaches r_f, as well as the time of flight Δt_f and the azimuthal variation $\Delta\varphi_\mathrm{f}$.

> Actually, `geokerr` works with the variables $1/r$ and $\cos\theta$ rather than r and θ, but this does not change the discussion.

14.1.5 Observables

For $r_\mathrm{f} \gg r_\mathrm{g}$, as said earlier, the outgoing direction α of the photon should be close to θ_f. However, since r_f is finite, it is slightly more accurate to compute α as

$$\cos\alpha = \boldsymbol{u}\cdot\boldsymbol{e}_z = \cos\theta_\mathrm{f} u^{\hat{r}} - \sin\theta_\mathrm{f} u^{\hat{\theta}}, \tag{14.17}$$

where $\boldsymbol{e}_z$ is the unit vector along the spin axis of the black hole. As expected, increasing r_f makes α tend to θ_f. Photons are then sent to different bins according to their outgoing direction α. We use an angular resolution of $\Delta\alpha = \pi/24$. Eventually, images and light curves can be reconstructed. The impact parameters of each photon can be computed according to Eq. (14.13) and (14.14), before stacking all the photons on the observer's screen to build an image. The screen resolution is $\Delta X = \Delta Y = 0.05\, r_\mathrm{g}$. Since the simulation is axisymmetric, we collect all photons regardless of their outgoing azimuthal velocity.

To build light curves, one must know the time delay from the emission of the photon to the observer. This delay includes the time Δt_f that it takes to reach r_out, and a geometrical delay δt from $r = r_\mathrm{out}$ to the observer's plane (see Fig. 14.2). Let us define a Cartesian basis $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$, where $\boldsymbol{e}_z$ is aligned with the spin axis of the black hole and where φ is measured in the $(\boldsymbol{e}_x, \boldsymbol{e}_y)$ plane with respect to $\boldsymbol{e}_x$. In this basis, the final outgoing direction of a photon is defined by its co-latitude α and its azimuth ψ: $\cos\alpha = \boldsymbol{u}\cdot\boldsymbol{e}_z$ and $\cos\psi = \boldsymbol{u}\cdot\boldsymbol{e}_x$ (see Fig. 14.2). This time delay is computed as (Cerutti et al., 2016)

$$c\,\delta t = r_\mathrm{out} - r_\mathrm{f}\left(\sin\alpha\sin\theta_\mathrm{f}\cos\left(\varphi_0 + \Delta\varphi_\mathrm{f} - \psi\right) + \cos\alpha\cos\theta_\mathrm{f}\right), \tag{14.18}$$

as can be seen on the geometrical construction in Fig. 14.2. Note that even if we have a "ring" of observers at fixed angle α, and all photons are collected regardless of ψ, the angle ψ intervenes because of the finite propagation time from the sphere at r_out to the observer's plane. The total time delay, from emission to collection by the observer, is $t_\mathrm{delay} = \Delta t_\mathrm{f} + \delta t$.

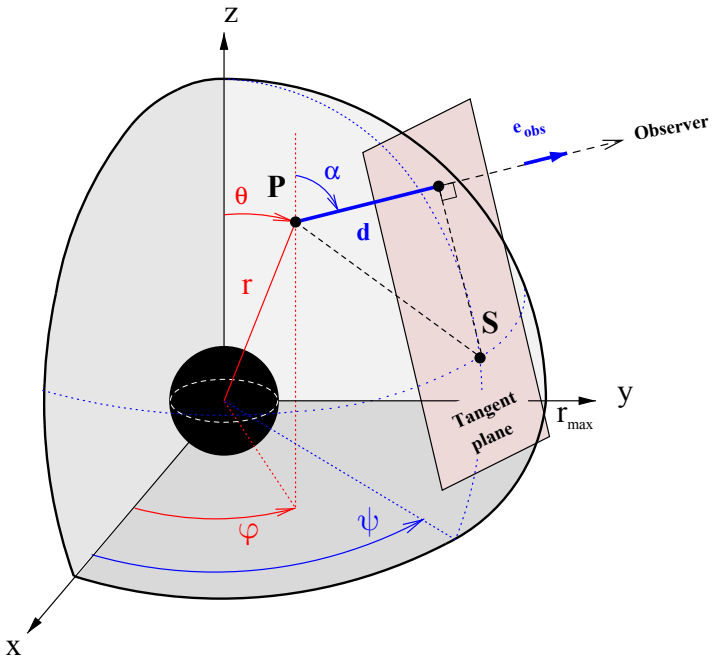


FIGURE 14.2: A photon is emitted at P $(r_\mathrm{f}, \theta_\mathrm{f}, \varphi_\mathrm{f})$ and propagates in the direction $\boldsymbol{e}_\mathrm{obs}$ defined by the angles (α, ψ). The observer is located at S, the intersection between the spherical screen and $\boldsymbol{e}_\mathrm{obs}$. The time delay is given by $\delta t = \boldsymbol{PS}\cdot\boldsymbol{e}_\mathrm{obs}/c$. Figure adapted from Cerutti et al. (2016).

All in all, photons observed at infinity with an angle α are staggered on a time grid according to their time of arrival. We use a time resolution of $\Delta t = 0.02\, r_\mathrm{g}/c$. Each photon contributes to the luminosity per unit solid angle as

$$\frac{\mathrm{d}L}{\mathrm{d}\Omega} = \frac{E}{2\pi \sin\alpha\, \Delta\alpha \Delta t}. \tag{14.19}$$

We do not keep any spectral information in the gamma-ray emission because of the low separation of scales. Our light curves are integrated over all photon energies. The information of around 10^4 photons per time step is saved.

Because `geokerr` was designed for backward ray tracing, a few adjustments had to be made to the code, allowing it to correctly pick the right geodesic given the photon parameters listed above. I verified that the two codes matched to high accuracy, regardless of the complexity of the geodesic (number of turning points in r or θ) or the number of points computed on the geodesic by `geokerr`. A typical photon trajectory is shown in Fig. 14.3(a), where the dashed line represents the trajectory of a photon computed by `Zeltron`, whereas the yellow solid line is an output from `geokerr` (for which a high number of points along the geodesic were computed to allow comparison). While `geokerr` conserves the constants of the motion to machine precision by construction, `Zeltron` does not. The energy-at-infinity, angular momentum and Carter constant along the `Zeltron` geodesic are shown as a function of the integration time pictured in Fig. 14.3(b), for the specific photon of Fig. 14.3(a). In general, the `Zeltron` particle pusher conserves E and Q to a relative accuracy of 10^{-8} at worst, which is very satisfying. Note that L is exactly conserved, because the simulation is axisymmetric: p_φ is not evolved by the PIC code.

14.2 Light curves

14.2.1 Results

As we saw in Chap. 13, the equatorial current sheet is prone to magnetic reconnection. This process has been invoked to be responsible for the very rapid flares in blazars or pulsar

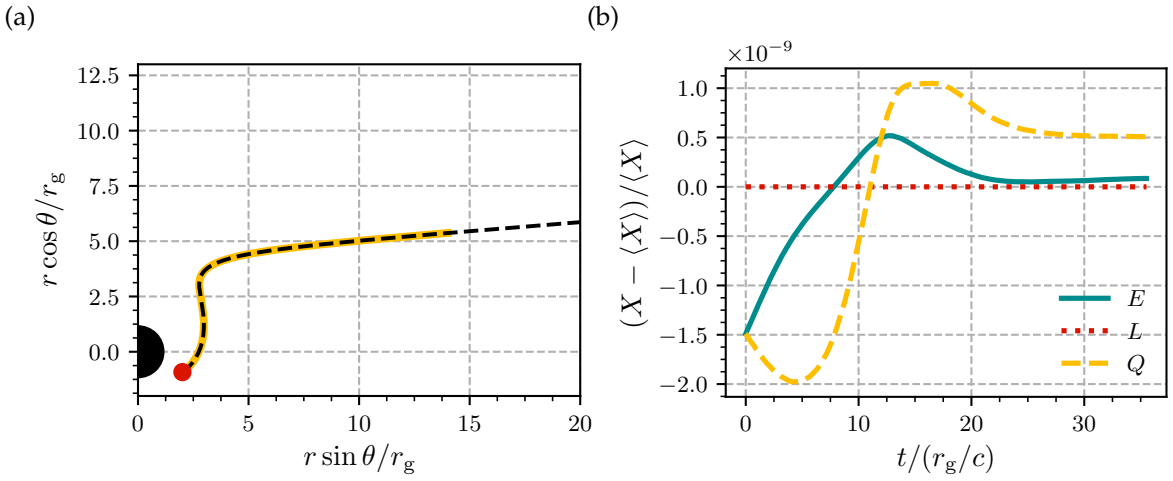


FIGURE 14.3: (a) Trajectory of a photon evolved by `Zeltron` (yellow solid line), compared with the reconstruction by `geokerr` using the photon's initial parameters (black dashed line). (b) Time evolution of the conserved quantities E, L and Q for this photon, as evolved by `Zeltron`.

wind nebulæ (Cerutti et al., 2012; Christie et al., 2019; Mehlhaff et al., 2020). It is intermittent and can accelerate particles very efficiently. Therefore, one might expect the dynamic current sheet from the paraboloid simulations in Chap. 13 to produce flares of VHE emission. This is also suggested by the strong variations in the magnetic flux and electromagnetic luminosity (see Fig. 13.8). It is possible to assess this scenario by investigating the synthetic light curves from these simulations, assimilating VHE emission to inverse Compton emission. We applied the previously outlined procedure to our three simulations with $V_0/c = 0.05$, along with the highest opacity monopole simulation presented in Chap. 12 (which had $\tilde{B}_0 = 5 \times 10^5$, $\tilde{\epsilon}_0 = 5 \times 10^{-3}$ and $\tau_0 = 30$).

The time-averaged luminosity corresponds to the total energy extracted from the black hole by the Blandford-Znajek process, while a fraction of it is channeled into gamma-ray photons. We define the time-averaged luminosity $\langle L_{\rm EM} \rangle$, and compute the energy flux per unit of solid angle using the method outlined in Sec. 14.1.5. The angle-resolved light curves are normalized by $\langle L_{\rm EM} \rangle / \Delta\Omega_{\rm obs}$, where $\Delta\Omega_{\rm obs} = 2\pi \sin\alpha \Delta\alpha$ is the solid angle in which originate the photons contributing to the light curve at viewing angle α.

Light curves extracted from the monopole simulation at $\tau_0 = 30$ are shown in Fig. 14.4(a). There is only a weak dependence on the viewing angle, especially at intermediate latitudes. This is expected, since the monopole simulations show little structure in the orthoradial direction, and photons are mainly emitted radially by particles flowing along the magnetic field lines. The bolometric inverse Compton luminosity of this simulation is $\langle L_\gamma \rangle = \langle \int (\mathrm{d}L_\gamma / \mathrm{d}\Omega)\, \mathrm{d}\Omega \rangle \simeq 0.04 \langle L_{\rm EM} \rangle$. This is consistent with the dissipation rate of electromagnetic energy measured in Sec. 12.2.4, confirming that the dissipated Poynting flux is mainly transferred to high-energy photons below the pair creation threshold. Although the light curve shows signs of rapid variability, which is consistent with the small size of the gap, it exhibits no flare. The incoherent process of pair creation along various magnetic field lines hinders the occurrence of large amplitude flares.

The situation is rather different for the paraboloid simulations with external forcing (Fig. 14.4(b), 14.4(c) and 14.4(d)). These light curves show pronounced differences if viewed face-on (line of sight close to the spin axis, low α) or edge-on (line of sight close to the equatorial plane, $\alpha \simeq \pi/2$). At low α, they exhibit stronger variability. During a "flaring" event, the flux doubles within a rising time $\simeq 2\, r_{\rm g}/c$. The periodicity of these flares is around $10\, r_{\rm g}/c$, in agreement with the periodicity of the giant plasmoid accretion cycles. Conversely, light curves observed at α close to $\pi/2$ are remarkably smooth, with no short-scale variability. In order to understand these qualitative differences, we have constructed two light curves, associated with the sites of emission of the photons. We have distinguished the polar cap as the zone defined by $\theta \in [0, \pi/3] \cup [2\pi/3, \pi]$, and the current sheet, defined by $\theta \in [\pi/3, 2\pi/3]$ and $r \in [r_{\rm h}, 5\, r_{\rm g}]$ (see the sketch in Fig. 14.5). The distinct light curves of these two zones are shown in Fig. 14.6. Unsurprisingly, photons emitted in the polar cap mainly contribute to the emission at low viewing angles, whereas the emission at $\alpha \simeq \pi/2$ is mainly due to the current sheet.

The fiducial optical depth τ_0 mainly impacts the intensity of the current sheet emission. In the simulations with external forcing, the average bolometric luminosity $\langle L_\gamma \rangle$ goes from $0.1 \langle L_{\rm EM} \rangle$ at $\tau_0 = 30$ to nearly $0.5 \langle L_{\rm EM} \rangle$ for $\tau_0 = 80$. Therefore, a very significant fraction of the electromagnetic power extracted from the black hole is converted to gamma-ray luminosity. The radiative efficiency is much higher than in the monopole simulations. Importantly, regardless of τ_0, the bolometric luminosity of photons emitted in the polar cap remains around 5% of the Blandford-Znajek luminosity. This fraction is similar to the monopole total high-energy bolometric luminosity, although the polar cap emission is much more variable in the paraboloid simulations. The excess luminosity entirely originates from the current sheet, which dissipates around 30% to 40% of the electromagnetic luminosity.

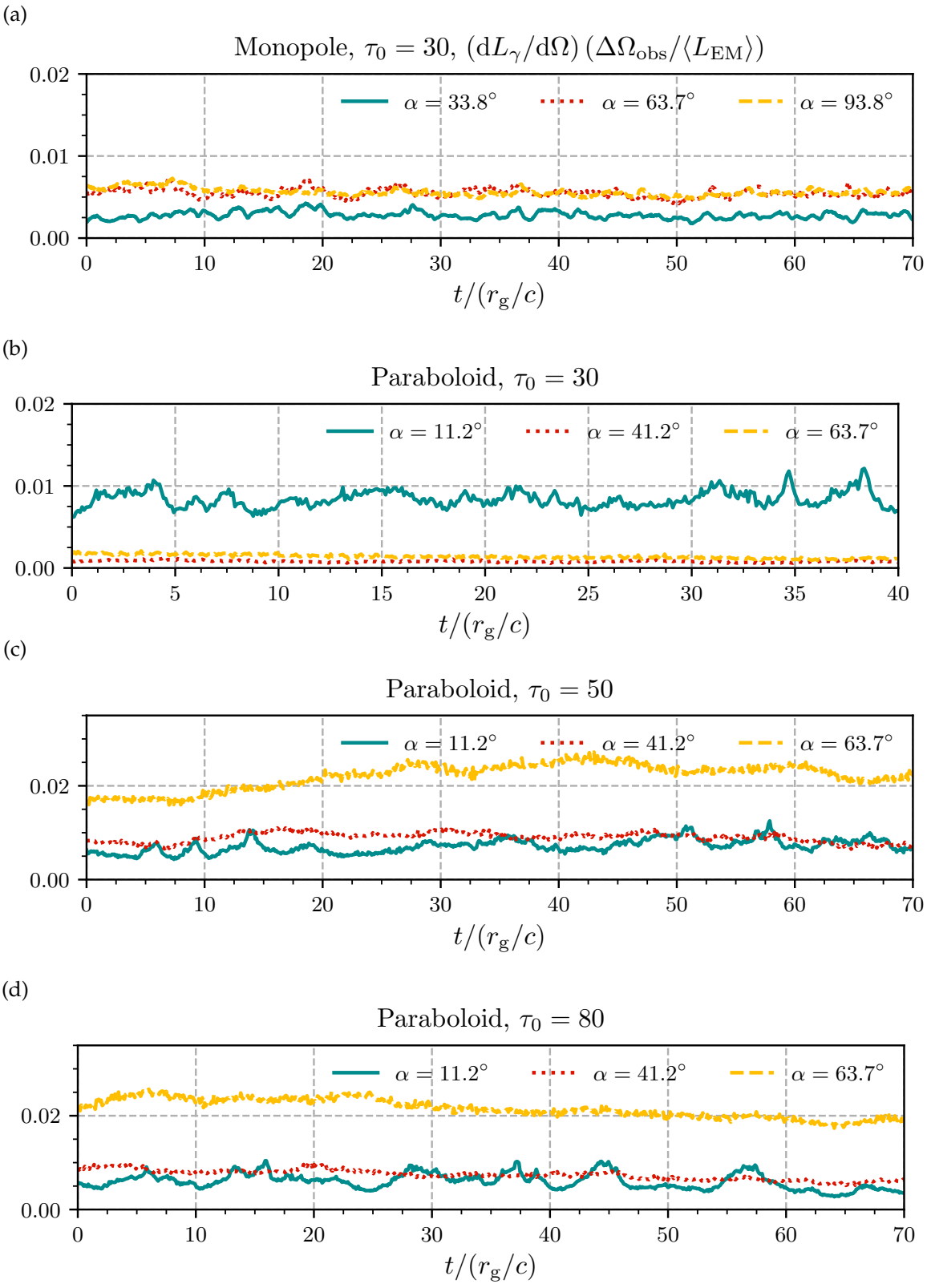


FIGURE 14.4: Light curves for the monopole simulation at $\tau_0 = 30$ (a), and for the paraboloid simulations with $V_0/c = 0.05$ and $\tau_0 = 30$ (b), $\tau_0 = 50$ (c) and $\tau_0 = 80$ (d). The blue solid lines denote low viewing angles, with the black hole seen face-on. The yellow dashed lines denote high viewing angles, with the black hole seen edge-on.

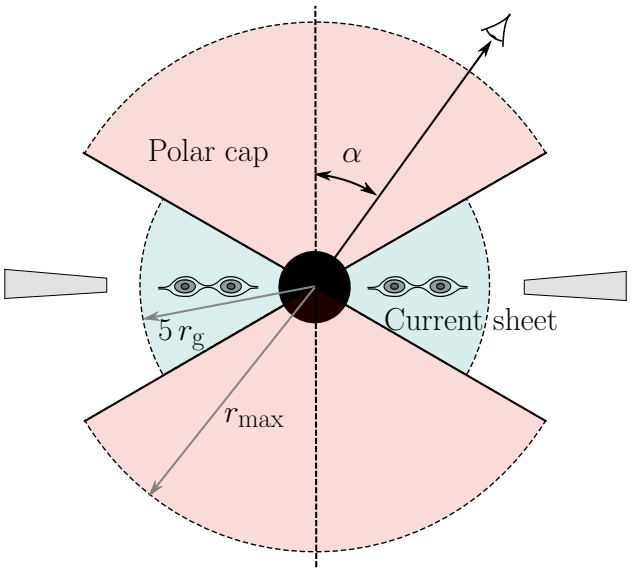


FIGURE 14.5: Sketch of the "polar cap" and "current sheet" zones. The gray shaded rectangles represent the conductive disk.

We can conclude unambiguously that high variability should not be expected from magnetospheres observed edge-on. This could stem from the fact that the formation of plasmoids in the current sheet, and the subsequent emission of high-energy photons, is inherently incoherent. However, several local studies of relativistic magnetic reconnection in a Harris sheet setup have posited that this process could induce strong flares. On the other hand, photons emitted in this region travel along complex null geodesics, which can have several turning points in θ. These geodesics are likely to differ significantly from simple radial rays; this makes the current sheet emission lose time coherence, which smears out any strong variability. Photons emitted from the polar cap follow more direct geodesics toward the observer at infinity, such that the variability of the primary process is imprinted in the light curve. Also, it should be noted that the polar cap shows more pronounced variability in these simulations than in the monopole case. This indicates that the gap dynamics cannot be studied with no consideration for the global magnetospheric structure: the magnetospheric

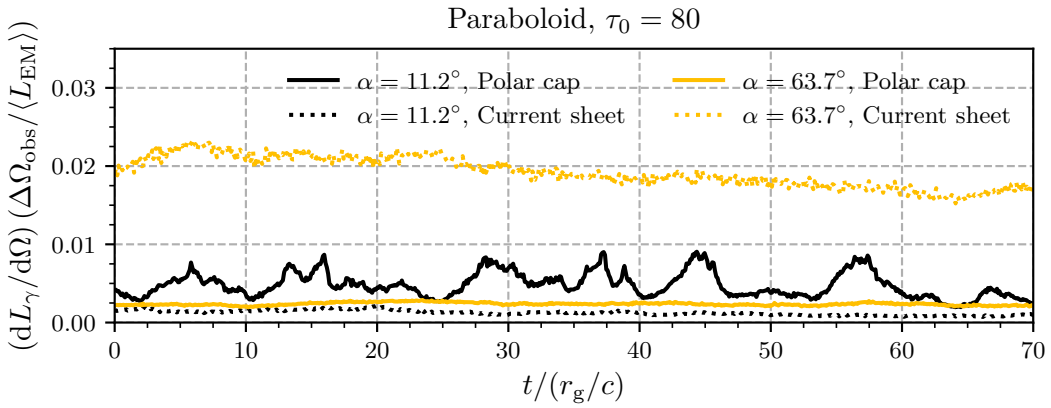


FIGURE 14.6: Light curve of the polar cap and current sheet in the paraboloid simulation with $V_0/c = 0.05$ and $\tau_0 = 80$, for a low and high viewing angles. The dotted lines represent the current sheet emission, whereas the solid lines represent the polar cap emission.

dynamics enhance the activity of the gap.

The power spectral densities of the light curves at $\alpha = 11.2°$ for the simulations with $V_0/c = 0.05$, computed using the `stingray` package (Huppenkothen et al., 2019), are shown in Fig. 14.7 (after logarithmic rebinning). At frequencies $0.1\,c/r_g \lesssim f \lesssim 2\,c/r_g$, it is well described by a red-noise power law $\propto f^{-p}$, with $p = 2.00 \pm 0.13$. A spectral break is visible around $0.1\,c/r_g$. The spectral break frequency is consistent with the characteristic timescale associated with giant plasmoid accretion events. Resolving this spectral break observationally would require data acquisition for much longer than $10\,r_g/c$ (more than 4 days in the case of M87*). Beyond $2\,c/r_g$, the power spectra are rather similar to a white noise. Most of the power is distributed at lower frequencies: flux variations on long timescales dominate those on short timescales. The value of the index p is in agreement with that measured by Aharonian et al. (2007) from the AGN PKS 2155-304, although this measure may depend on whether the AGN is in a flaring state or not (H. E. S. S. Collaboration et al., 2017). We find that the value of p does not depend much on τ_0. The characteristic value of the plasma frequency $\nu_p = \omega_p/2\pi = \sqrt{8\pi e^2 \langle n/\Gamma \rangle / m_e}/2\pi$ is $50\,c/r_g$, and lies beyond the frequency range presented on the spectrum.

14.2.2 Discussion

We find that a substantial fraction of the Blandford-Znajek electromagnetic luminosity is channeled into inverse Compton photons. This high radiative efficiency is the most salient feature of these paraboloid simulations, especially in the presence of external forcing $V_0 \neq 0$. It confirms the view adopted in Chap. 12 that the dissipated electromagnetic energy was largely transmitted to gamma-ray photons. We emphasize that emission from equatorial latitudes is smoothed out, so that high-energy variability should primarily be expected from the polar caps. Variability from the polar caps is enhanced with respect to the monopole simulations. We also observe that in the absence of magnetic flux supply, the magnetosphere reaches a somewhat quiescent state and cannot produce flares. This is consistent with the recent resistive GRMHD simulations of Ripperda et al. (2020), who found that only in the magnetically arrested disk setup could there be a flaring state, during which plasmoids formed in an equatorial current sheet are heated to relativistic temperatures.

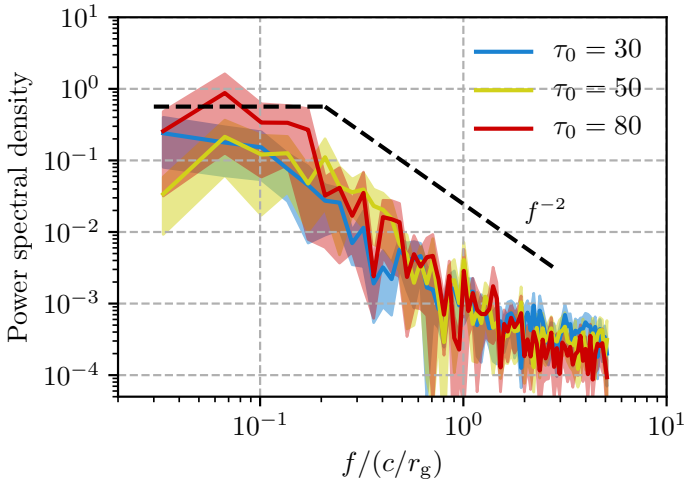


FIGURE 14.7: Power spectral density of the light curves at viewing angle $\alpha_{obs} = 11.2°$, for the simulations with $V_0/c = 0.05$, as a function of the frequency f. The black dashed line shows the red-noise scaling f^{-2}, down to a spectral break located close to $0.1\,c/r_g$. The shaded areas indicate the error bars on the power spectra.

However, even with constant external forcing, the intrinsic activity of a steady-state black-hole magnetosphere does not reproduce the most dramatic features of AGN flares: a flux-doubling time below $r_{\rm g}/c$ and an increase in the flux by a factor of at least 5, rather than 2 in our modeling. The variability seems to be well characterized by a red-noise power law, down to $\simeq 2\,c/r_{\rm g}$. Because of the nonaxisymmetric modes mentioned earlier, 3D simulations are likely to show even less variability. It should also be pointed that in our simulations, the background radiation field is mono-energetic. Using a more realistic power law would have reduced the variability in the gap screening, because the pair creation threshold would not have been defined at a single photon energy. This makes it even less likely that realistic gamma-ray flares can be accommodated with our numerical setup.

If of magnetospheric origin, rather being a manifestation of the intrinsic variability due to the pair production mechanism, flares could be interpreted as the fast response of a black-hole magnetosphere to a sudden change in the external parameters. This conclusion was also reached by Levinson & Cerutti (2018) and Kisaka et al. (2020), through radiative 1D GRPIC simulations. For example, a variation in the accretion rate would cause the density of soft photons to increase, leading to an increase in τ_0. The velocity of magnetic field transport could also change, should a very magnetized plasma blob accrete toward the magnetosphere. In that sense, black-hole magnetospheres differ fundamentally from pulsar magnetospheres: pulsar activity is determined by parameters that are characteristic to the pulsar itself, and therefore remains quite stable.

14.3 Images

14.3.1 Synchrotron radiation

Very much like we used inverse Compton emission as a proxy for VHE radiation, we assume that millimeter wavelengths are adequately modeled by optically thin synchrotron emission. Quantum effects are neglected, so that we can work in the approximation where synchrotron photons are emitted continuously by gyrating particles, with energies much smaller than their parent particle's energy. Also, unlike gamma-ray photons, we do not track synchrotron photons in the PIC simulations. Rather, at each time step, we compute the parameters of the photons emitted by a fixed fraction $f = 1\%$ of all electrons and positrons, and dump them to an external file for later `geokerr` treatment.

At each time step, a macro-photon emitted by a charged particle carries the total power $P_{\rm tot}$ emitted by the particle, which is given by Eq. (8.43). We make the delta approximation, approximating the power spectrum of the macro-photon with $P_{\rm tot}\delta\left(\nu-\nu_{\rm s}\right)$, where $\nu_{\rm s}$ is the characteristic synchrotron frequency (see Eq. (8.50)). This approximation is valid if the spread in the emitting particle spectrum is much larger than the width of the synchrotron kernel, given in Eq. (8.50).

The synchrotron photons are assumed to be emitted along their parent particle's momentum, in the ultra-relativistic limit. First, the effective perpendicular magnetic field (see Eq. (8.44)) is computed at the emitting particle's location. It is used to compute the emitted frequency $\nu_{\rm s}$ and the power $P_{\rm tot}$ of the macro-photon. The photon frequency measured by an observer at infinity ν_∞ is related to the emitted frequency $\nu_{\rm s}$ as in Eq. (14.16). It is an output parameter dumped in a text file, but it is not used by `geokerr`. The synchrotron power is encoded in the weight of the photon. A single time-averaged image, for a chosen viewing angle, is made of around 10^7 photons.

14.3.2 Test cases

We tested the image reconstruction procedure against several test cases. Our goal is to numerically retrieve the black-hole shadow. If a Kerr black hole is backlit by a distant source of light of angular size much larger than the black hole, it will cast a shadow on the screen of the observer. The interior of the shadow corresponds to unstable orbits which end up in the black hole, resulting in a dark region on the observer's screen. The size of the shadow is larger than $r_{\rm h}$ because of strong light bending. The rim of the shadow on the screen is hit by photons which have circled the black hole many times, on bound, marginally stable orbits. An implicit expression for the rim of the black-hole shadow and a physical interpretation its the dependence with a can be found in the Appendix 14.A. This special curve was alternatively dubbed the "critical curve" by Gralla et al. (2019).

In the case of a Schwarzschild black hole, the shape of the shadow does not depend on the viewing angle. It is a circle of radius $r_{\rm ph} = \sqrt{27}\,r_{\rm g}$. As shown by Luminet (1979), if the black hole is backlit by a plane-parallel beam of light, the image presents a dark disk of radius $r_{\rm ph}$. $r_{\rm ph}$ is the critical impact parameter below which back-traced photons would be captured by the black hole. Gralla et al. (2019) noted that the image of a black hole lit by a distant spherical screen would present the same feature. Falcke et al. (2000) and Narayan et al. (2019) have designed semi-analytical models to compute the image of a black hole surrounded by a spherically symmetrical, optically thin accretion flow, up to the event horizon. This case is less straightforward, because some photons can have impact parameters smaller than $r_{\rm ph}$ but still escape to infinity, since they are not emitted infinitely far from the black hole. Hence, the "shadow" is not exactly dark. However, the region inside $r_{\rm ph}$ still shows a deficit in brightness, because a larger fraction of photons is captured by the black hole. Besides, those photons experience a stronger redshift, leading to a lower contribution to the intensity on the image. For spherically symmetric accretion flows, the image is quite insensitive to the details of the plasma emission properties, and matches the analytical shadow.

In our setup, the simplest test case is to embed the black hole in a spherically symmetric region populated by synchrotron photons, from $r_{\rm h}$ to $10\,r_{\rm g}$. The radiation field is uniform, isotropic and mono-energetic. An example of resulting image is shown in Fig. 14.8(a). The analytical prediction for the rim of the shadow is shown as a dashed line. For all viewing angles, we find a circular dark region centered on $(X = 0, Y = 0)$. Fig. 14.8(b) shows the radial intensity profile, averaged on the viewing angle α. The intensity peaks close to the prediction $r_{\rm ph} \simeq 5.2\,r_{\rm g}$. As noticed earlier, the intensity is not zero inside the shadow. The shape of the intensity profile is similar to that found by Narayan et al. (2019), although the emissivity profile is different in their case.

The reasoning is similar for a Kerr black hole. Since the black hole will capture most photons orbiting within the marginally stable circular orbits, if it is embedded in a spherically symmetric accretion flow, its image should show a brightness depression consistent with the shadow (Falcke et al., 2000). As can be seen in Fig. 14.8(c) and 14.8(d), we indeed observe such a shadow. Its rim is consistent with the analytical prediction. Note that while the shadow is shifted to the right on the X axis, with respect to the black hole's direction, the diameter of the shadow barely depends on a (Johannsen & Psaltis, 2010), making the diameter of the shadow a reliable marker of the black hole's mass. The emission on the left of the image is brighter because of Doppler beaming.

14.3.3 Results

In more complex configurations than an spherical accretion flow, the image does not necessarily display a clean shadow. Luminet (1979) constructed, for the first time, the image of a geometrically thin and optically thick accretion disk around a Schwarzschild black hole. More systematic studies have been performed to study the looks of a thin disk (Gralla et al.,

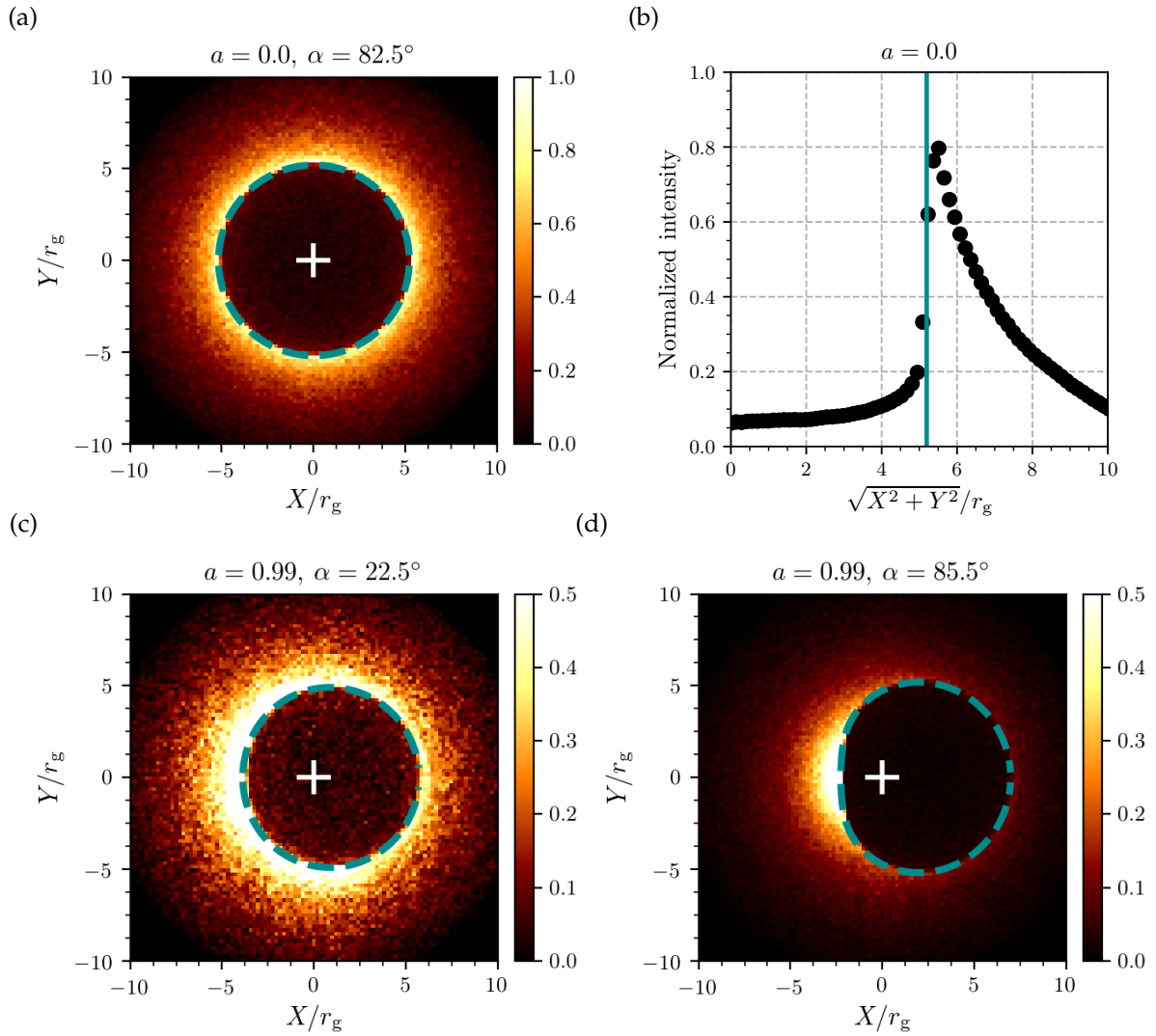


FIGURE 14.8: (a) Image of a Schwarzschild black hole with a viewing angle $\theta_0 = 82.5°$. (b) Radial intensity profile of the black-hole shadow with $a = 0$, averaged over θ_0, for the images at $a = 0$. (c) Image of a Kerr black hole with $a = 0.99$ with a viewing angle $\theta_0 = 22.5°$. (d) Same as (c) with $\theta_0 = 85.5°$. The intensity is in arbitrary units. The blue dashed curve denotes the analytical rim of the shadow.

2019), a moderately thick disk (Vincent et al., 2021) or a conical jet (Vincent et al., 2019). Bronzwaer et al. (2021) studies in detail which conditions must be met by the black-hole environment for the image to display the analytically predicted shadow.

On the other hand, no study has so far tried to model the contribution of the inner magnetosphere, especially from *ab initio* simulations. I have applied the image reconstruction technique described in the previous section to the simulation with $\tau_0 = 50$ and $V_0/c = 0.05$, presented in Chap. 13. The images depend little on the observing frequency, except for frequencies much smaller than $\nu_c = eB_0/m_e c$, at which smaller magnetic fields are probed, resulting in a much larger and more diffuse emission zone. We are interested in optically thin emission, which generally occurs at frequencies larger than the cyclotron frequency. In the case of the EHT for example,we have $\nu_c \simeq 1\,\mathrm{GHz}$, which is much smaller than the observing frequency 230 GHz. As a result, we will study images averaged in frequency over a bandwidth $[0.1\nu_c, 10\nu_c]$. We will also focus on time-averaged images. The resulting images are shown for various viewing angles in Fig. 14.9. Several features can be distinguished.

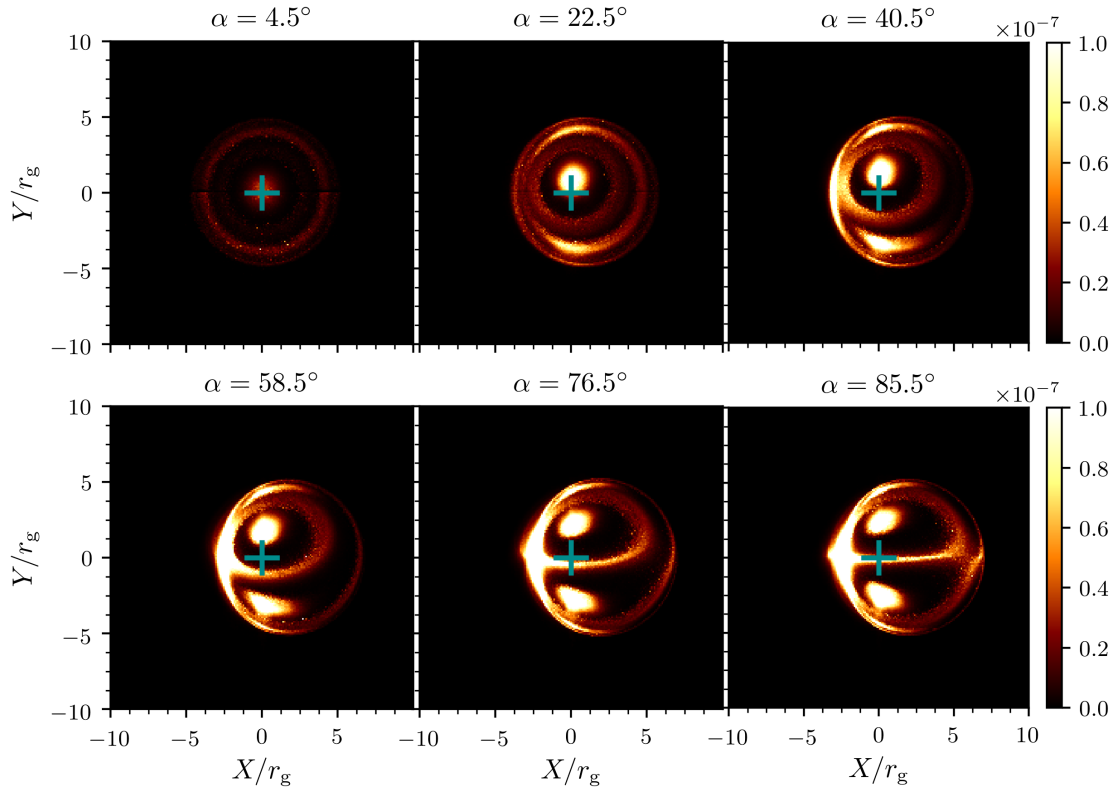


FIGURE 14.9: Time-averaged images seen at various viewing angles θ_0 from 0 to $\pi/2$ for the paraboloid simulations with $\tau_0 = 50$ and $V_0/c = 0.05$, in arbitrary units, averaged over $\nu \in [0.1\,\nu_c, 10\,\nu_c]$. The blue cross indicates the center of the image.

There is a central hot spot close to the center of the image at low viewing angles, which turns into two bright zones symmetric with respect to the X axis at larger α. These bright spots are surrounded by several lensed rings. In order to determine the physical origin of these different components, we applied the same decomposition as in Sec. 14.2, separating the emission from the current sheet and the polar cap. These two emission zones were sketched in Fig. 14.5.

The distinct emission from the current sheet is shown in Fig. 14.10(a), whereas that from the polar cap is shown in Fig. 14.10(b). This decomposition separates unambiguously the different components in the image. The central hot spot at low viewing angles originates solely from the polar cap. It is the direct image of the upper photons emitted in the polar cap by pair discharges. The surrounding ring is the direct image from the lower polar cap, on the opposite side of the black hole, which is strongly lensed. We checked that the photons from this ring crossed the equatorial plane once, whereas the ones from the hot spot did not. At larger α, the contributions from the two sides of the black hole tend to be axially symmetric.

The current sheet emission is negligible at very low α. At intermediate viewing angles, the image displays a central moderately thick ring and a wider thin ring. The rings' brightness is larger on the left side of the image, as a result of Doppler beaming. The inner ring is the direct image of the radiating current sheet, whereas the outer ring is the secondary image which has been lensed by the black hole. The shape and location of the outer ring match the analytical rim of the shadow. Close to an edge-on view ($\alpha \simeq \pi/2$), the current sheet can be directly seen from the edge as emission on the X axis, and indirectly in a surrounding ring. This is not unlike the image of an optically thin accretion disk (Luminet, 1979).

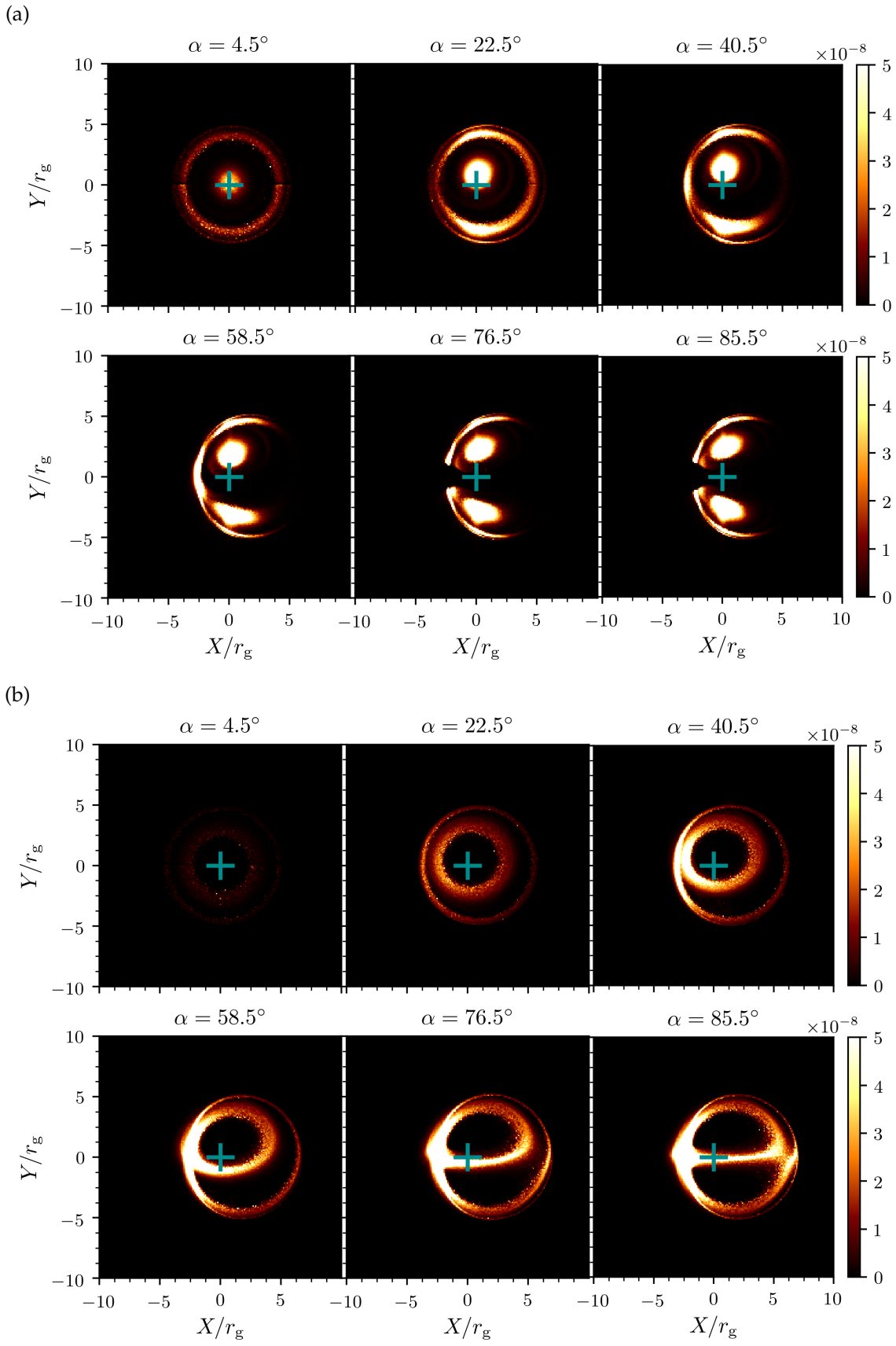


FIGURE 14.10: Contribution from the polar cap (a) and the current sheet (b), as defined in Fig. 14.5, to the image in Fig. 14.9.

14.3.4 Toy models

The morphologies of these images can help us constrain the physics of the emission. I built toy models of synchrotron emission around the black hole, which are represented in Fig. 14.11(a) and 14.11(b), with $a = 0.99$. Toy model 1 consists in polar cap emission from a zone delimited by $r \in [r_\mathrm{h}, 2.5\,r_\mathrm{g}]$ and $|\theta| \leq 0.4$, $|\pi - \theta| \leq 0.4$ (see Fig. 14.11(a)). Toy model 2 (resp. 3) models current sheet emission from a zone delimited by $|\pi/2 - \theta| \leq 0.05$ and $r \in [r_\mathrm{h}, 2\,r_\mathrm{g}]$ (resp. $r \in [r_\mathrm{h}, 6\,r_\mathrm{g}]$) (see Fig. 14.11(b)). We filled these zones with a uniform, isotropic and mono-energetic synchrotron radiation field, as in the test cases (Sec. 14.3.2). The images associated with model 1 are shown in Fig. 14.12(a), whereas the images from model 2 are shown in Fig. 14.12(b).

The images from toy model 1 are exactly consistent with the polar cap emission, confirming that the magnetosphere presents bright emission spots at the poles of the black hole. The case of the current sheet emission is more interesting. It is not clear *a priori* what the size and extent of the equatorial emission are. In the case of the toy model 2, where the emission extends to $2\,r_\mathrm{g}$ and is concentrated at the innermost regions, the images show distinctly one direct ring and one secondary ring (see Fig. 14.12(b)). This is precisely what is observed in Fig. 14.10(b). On the other hand, if the equatorial emission zone extends up to $4\,r_\mathrm{g}$ or $6\,r_\mathrm{g}$, the direct emission is more diffuse (see Fig. 14.13). The differences between the two models are most pronounced for intermediate viewing angles, at which the two rings can no longer be distinguished. At edge-on view or face-on views, their predictions are similar. Therefore, in our simulations, a major fraction of the current sheet image must originate from a very compact zone, probably within $2\,r_\mathrm{g}$. Another teaching from such toy models is that the equatorial origin of the emission ensures the existence of the central brightness depression. We checked that with a thicker emission zone, the inner dark region shrinks.

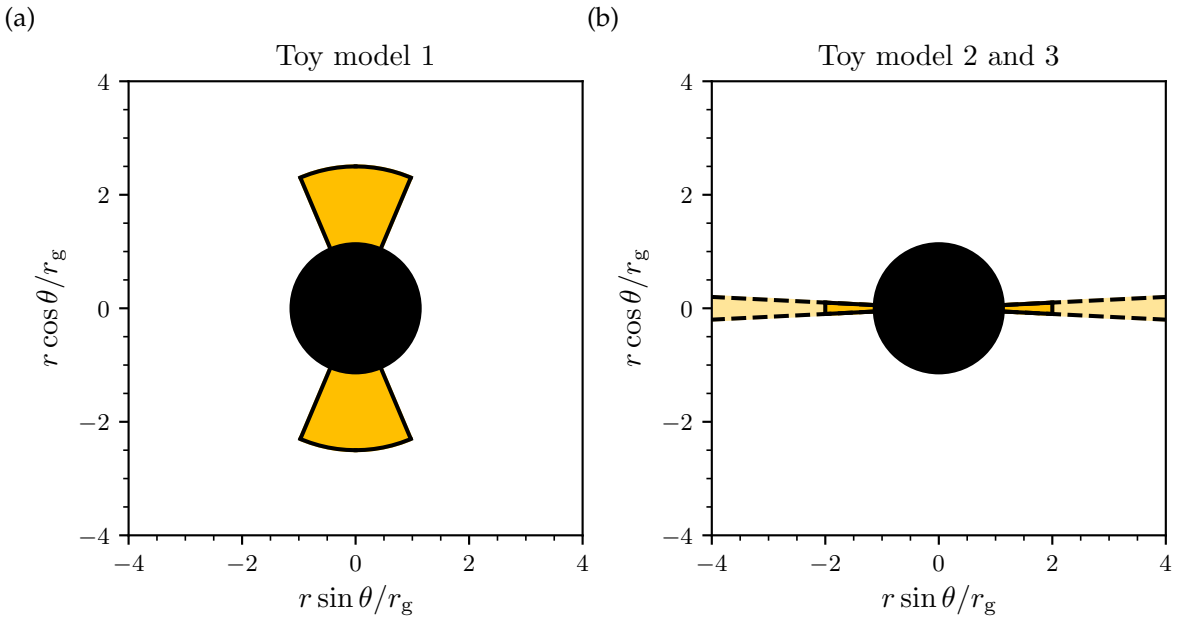


FIGURE 14.11: Sketch of the emitting zone in the different toy models (yellow area). (a) Polar cap emission model. (b) Current sheet emission model; in model 2 the emitting zone extends to $2\,r_\mathrm{g}$ (solid line), whereas it extends to $6\,r_\mathrm{g}$ in model 3 (dashed line).

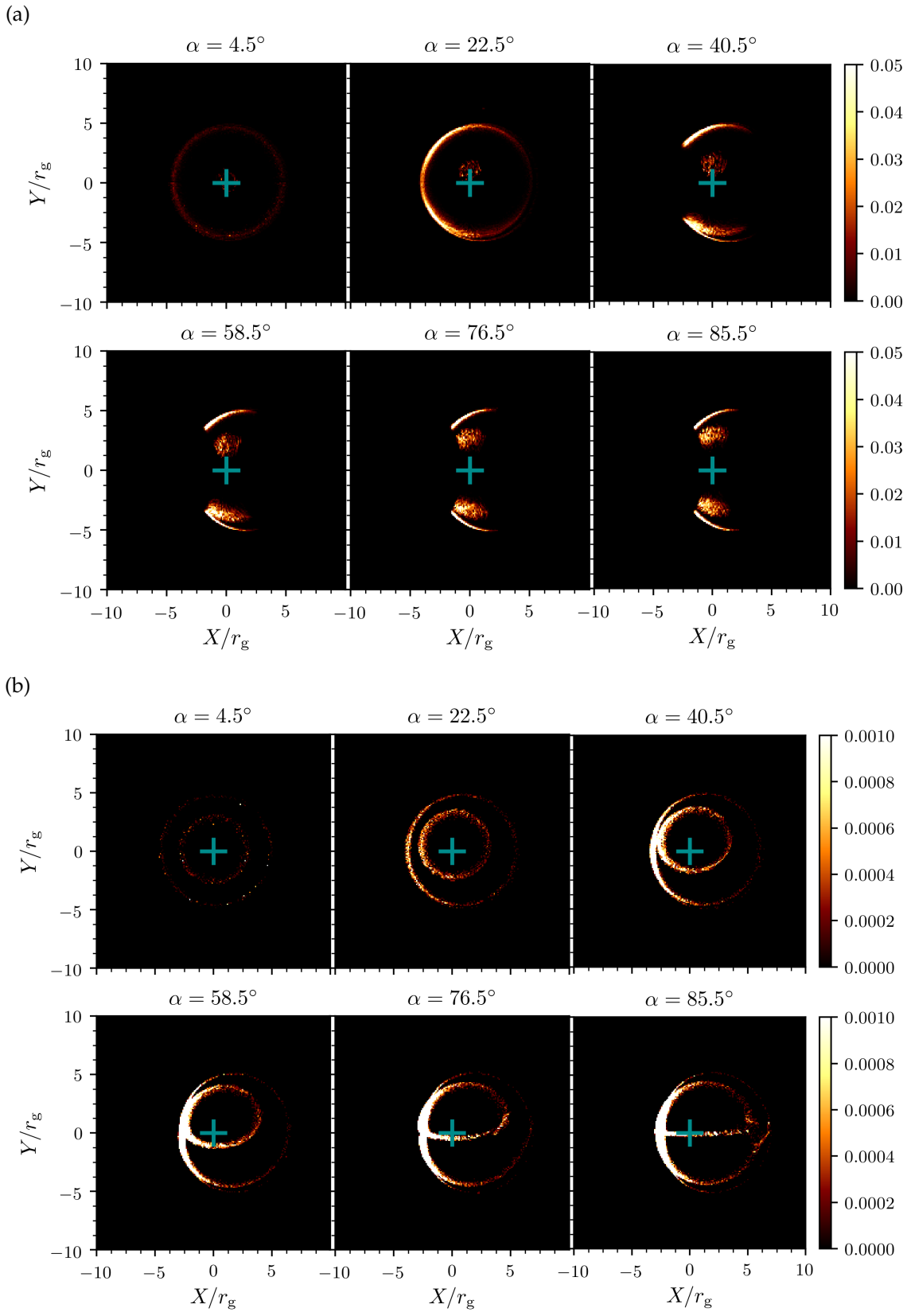


FIGURE 14.12: Images from the toy model 1 (a) and 2 (b) at various viewing angles, in arbitrary units. The blue cross indicates the center of the image.

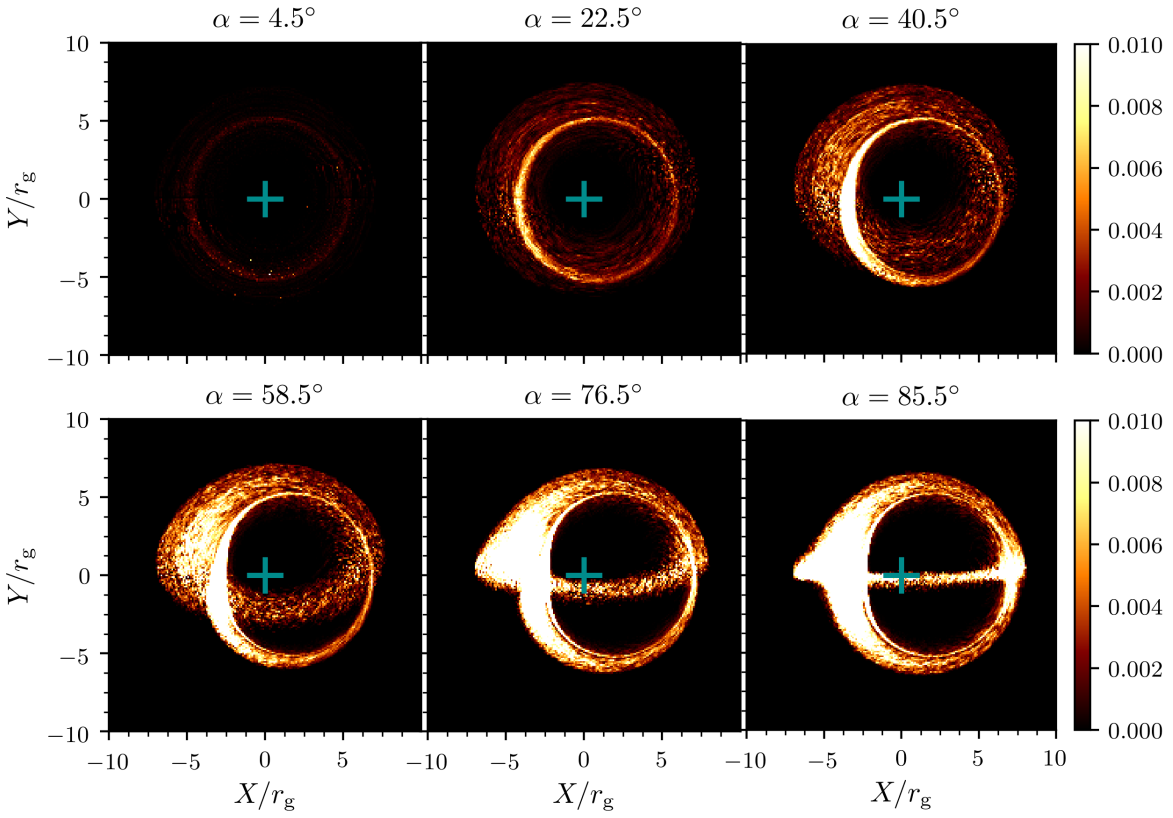


FIGURE 14.13: Same as Fig. 14.12 for the toy model 3, with a current sheet emission up to $6\,r_g$.

14.3.5 Interpretation and application to M87*

In order to apply these findings to the EHT images of M87*, let us assume that the spin of the black hole is aligned with the large-scale jet, so that it makes an angle $\alpha \simeq 18°$ with our line of sight (Walker et al., 2018). In the sixfold image panels presented so far, the EHT data should be compared to the top-middle images. The main discrepancy between the images produced by our simulations and the ones observed by the EHT is the central bright spot. However, it is possible that the intense synchrotron emission from the polar cap is overestimated in our simulations, due to the absence of synchrotron cooling. Particles accelerated at the X-points in the current sheet are propagating in a zone where the magnetic strength is very low by construction, implying that they can suffer only moderate synchrotron losses. In contrast, the magnetic strength is high at all times in the polar caps, so that accelerated particles should cool quickly. Therefore, we expect the current sheet emission to dominate the image in the presence of synchrotron cooling. It is one of my main upcoming objectives to find out whether this hypothesis is valid, by including a radiation reaction force in the simulation (see Sec. 8.3.4).

Let us make the assumption that the main contribution to the image seen at $\alpha = 18°$ comes from the current sheet. From this magnetospheric model, we can make the prediction that when the EHT reaches a finer resolution, it will resolve two distinct rings. Fig. 14.14 shows two images of the current sheet emission from our simulation seen at $\alpha = 18°$, blurred with a Gaussian kernel with two different standard deviations σ (in the case of M87*, $5.2\,r_g \simeq 20\,\mu as$). The brightness asymmetry and morphology of the blurred images are consistent with the EHT observations. The originality of the magnetospheric model is to include equatorial emission up to the event horizon, whereas most models either assume

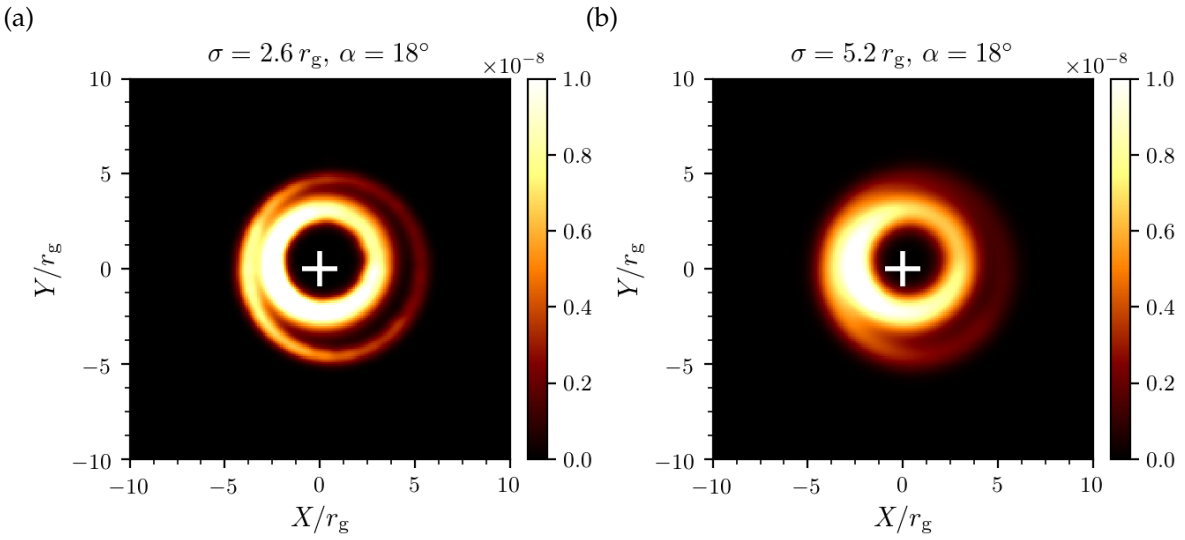


FIGURE 14.14: Image produced by the current sheet region from the paraboloid simulation with $\tau_0 = 50$ and $V_0/c = 0.05$, observed with $\alpha = 18°$, in arbitrary units, blurred with a Gaussian kernel at $\sigma = 2.6\,r_\mathrm{g}$ (a) and $5.2\,r_\mathrm{g}$ (b). The white cross indicates the center of the image.

that the emission zone is truncated at some larger radius, or that it is almost spherically symmetric. Instead, we find that if the polar caps are indeed quenched by synchrotron cooling, a compact and annular emission zone close to the event horizon can account for the EHT images at current angular resolution. In this case, the central brightness depression is well within the photon ring. If the current sheet emission brings the dominant contribution to the EHT image, this could readjust the estimation of the mass of M87*.

In this preliminary work, I have only discussed the morphologies of the synthetic images, without estimating the fraction of the electromagnetic luminosity that is actually converted into synchrotron radiation. It will be important to compare the synchrotron luminosity in my simulations to the brightness temperatures from the EHT observations, in order to test the magnetospheric scenario. This will also require the inclusion of synchrotron cooling in the code.

14.A Appendix: Black-hole shadow

In this appendix, we aim to derive an analytical expression for the rim of the shadow as a function of the spin parameter a and the viewing angle α. Let us trace back a photon with impact parameters (X, Y) to the source; it must have a radial turning point. A photon orbiting exactly at a radial turning point has $\dot{r} = 0$ and stays at constant t, on a "circular" orbit. What we must do is find the characteristics of the marginally stable circular orbits and relate them to the impact parameters. The properties of circular photon orbits in the Kerr metric are studied in detail by Bardeen et al. (1972) and Teo (2003). A construction of the shadow is provided by Frolov & Zelnikov (2011).

The marginally stable circular orbits correspond to the critical case, when two real roots of $\mathcal{R}$ coincide. Such an orbit with radius r_0 is defined by

$$\mathcal{R}(r_0) = 0, \quad \left.\frac{\mathrm{d}\mathcal{R}}{\mathrm{d}r}\right|_{r_0} = 0. \tag{14.20}$$

These two equations allow us to express ℓ and q^2 as a function of the radius of the orbit r_0, fully parametrizing the shadow on the observer's screen according to Eq. (14.13) and (14.14). They yield two classes of solutions:

$$\ell = \frac{\overline{r}^2 + a^2}{a}, \qquad q^2 = -\frac{\overline{r}^4}{a^2}; \qquad \text{(Class 1)} \qquad (14.21)$$

$$\ell = -\frac{\overline{r}^3 - 3\overline{r}^2 + a^2\overline{r} + a^2}{a\,(\overline{r}-1)}, \qquad q^2 = -\overline{r}^3\frac{\overline{r}^3 - 6\overline{r}^2 + 9\overline{r} - 4a^2}{a^2\,(\overline{r}-1)^2}; \qquad \text{(Class 2)} \qquad (14.22)$$

with $\overline{r} = r_0/r_\mathrm{g}$. Actually, the first class given by Eq. (14.21) is unphysical. For the geodesic to be physical, the angular potential Θ must be positive (see Eq. (14.3)). Here we have $q^2 < 0$, so that for $\Theta = q^2 + a^2\cos^2\theta - \ell^2\,\mathrm{cotan}^2\,\theta$ to be positive, the condition $a^2 - q^2 - \ell^2 > 0$ must be satisfied. However, inserting Eq. (14.21) into this condition leads to $a^2 - q^2 - \ell^2 = -2\overline{r}^2 < 0$. In the second class, one can check that $a^2 - q^2 - \ell^2 < 0$ in general, implying that it yields physical solutions only if $q^2 > 0$. Solving the cubic equation, one can show that q^2 is positive for $r \in [r_-, r_+]$, where

$$r_\pm = 2r_\mathrm{g}\left(1 + \cos\left(\frac{2}{3}\arccos\left(\pm|a|\right)\right)\right). \qquad (14.23)$$

All bound orbits lie at a fixed radius r_0, which must lie in $[r_-, r_+]$. Importantly, an orbit which remains in the equatorial plane $\theta = \pi/2$ must have $q^2 = 0$. Therefore, such equatorial circular orbits can only exist at $r_0 = r_+$ or $r_0 = r_-$. Although photon orbits around a Schwarzschild black hole are all confined to a plane passing through its center, it is not the case in the Kerr metric. There exists spherical orbits with constant radius r_0 which are not confined to the equatorial plane. However, such orbits are bound to remain in the radial shell $[r_-, r_+]$.

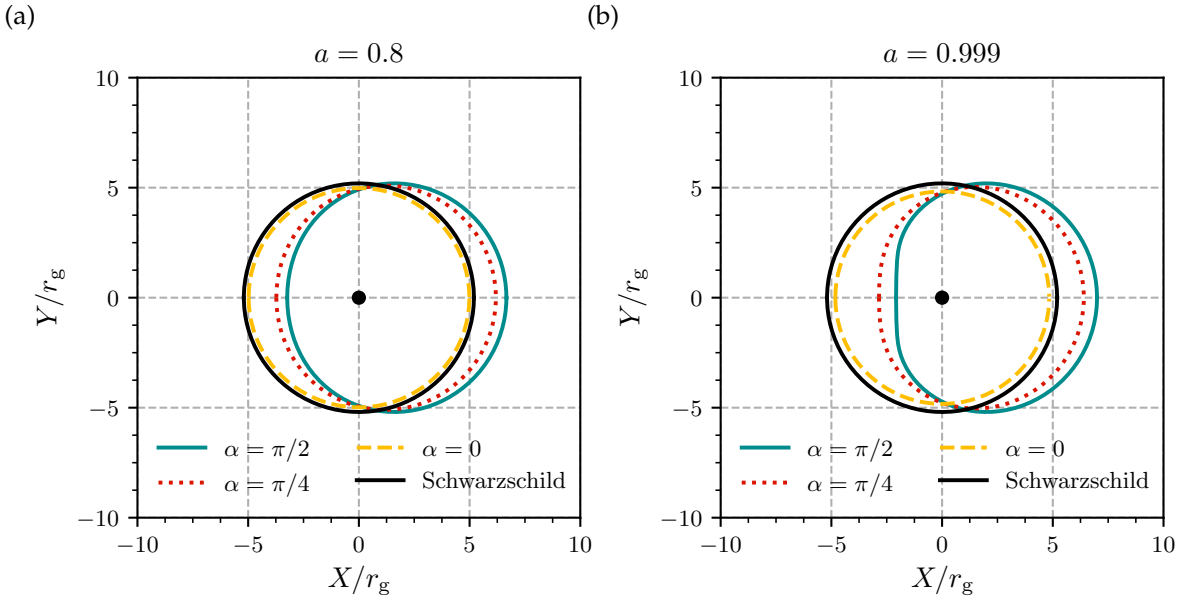


FIGURE 14.15: Rim of the shadow of a Kerr black hole with spin parameter $a = 0.8$ (a) and $a = 0.999$ (b) on the observer's screen with coordinates (X, Y), for several viewing angles α. The shadow of the Schwarzschild black hole (circle centered on $(0, 0)$ with radius $3\sqrt{3}\,r_\mathrm{g}$) is also shown (black solid line).

The rim of the shadow at observing angle α can be eventually be obtained by tracing the parametric curve

$$X(r_0) = -r_{\rm g}\frac{\ell(r_0)}{\sin\alpha}; \quad Y(r_0) = \pm r_{\rm g}\sqrt{q^2(r_0) + a^2\cos^2\alpha - \ell^2(r_0)\,\mathrm{cotan}^2\,\alpha}, \tag{14.24}$$

where $\ell(r_0)$ and $q^2(r_0)$ are given by Eq. (14.22), for $r_0 \in [r_-, r_+]$. Each r_0 parametrizes a circular (or rather, spherical) orbit of radius r_0. The resulting shadows for various angles and spin parameters are shown in Fig. 14.15.

In the limit $a = 0$, we find using Eq. (14.24) that $X^2 + Y^2 = r_{\rm g}^2\left(q^2 + \ell^2\right)$. Besides, in this case, the radial potential reads $\mathcal{R}(r_0) = \overline{r}^4 - \left(q^2 + \ell^2\right)\left(\overline{r}^2 - 2\overline{r}\right)$. Implementing the conditions given by Eq. (14.20) with this potential yields $r_0 = 3\,r_{\rm g}$ and $q^2 + \ell^2 = 27$. All in all, in the Schwarzschild limit, marginallly stable circular orbits all have radius $r_0 = 3\,r_{\rm g}$, and the resulting shadow is a circle radius $3\sqrt{3}r_{\rm g}$, regardless of the observer's inclination (see Sec. 2.2.4). The Schwarzschild shadow is also represented in Fig. 14.15 for comparison.

Let us interpret the shadow of a Kerr black hole close to maximal spin parameter (see Fig. 14.15(b)). Rays propagating in the $\alpha = 0$ direction have $\ell = 0$. This uniquely determines the radius of the spherical orbits which contribute to this image. Therefore, the shadow as seen by a face-on observer is a circle. As α gets closer to $\pi/2$, the circle shifts to the right and gets flattened at its left edge. The shift is largest for $\alpha = \pi/2$ and $a = 1$. A physical explanation of this shift is illustrated in Fig. 14.16. Corotating photon orbits are stabilized by frame-time dragging, so that the marginally stable circular photon orbits happen to be located closer to the black hole as a increases. The effect is opposite for counter-rotating orbits. Photons with positive impact parameters X have $\ell < 0$, so they were in counter-rotation with respect to the black hole, whereas photons with negative impact parameters were in corotation. Thus, the rim of the shadow is shifted to the right (in the X direction) for both corotating and counter-rotating photons.

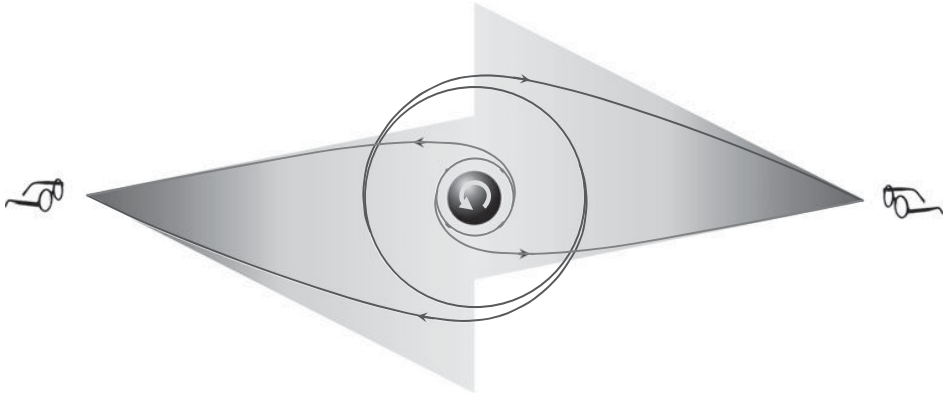


FIGURE 14.16: Photons hitting the left of the rim were propagating on corotating, close orbits, whereas photons hitting the right of the rim were propagating on counter-rotating, more distant orbits. Figure taken from Frolov & Zelnikov (2011).

Part V

Conclusion

"The only principle that does not inhibit progress is: anything goes."

"Le seul principe qui n'entrave pas le progrès est : tout est bon."

PAUL FEYERABEND, Against method

Chapter 15
Conclusion and perspectives

General relativistic kinetic simulations are a very recent tool, yet they have already offered new perspectives on the modeling of black-hole magnetospheres. Crucially, they can model properly the tenuous and magnetized funnel regions where plasma injection is supposed to take place. Using this tool, it is possible to include *ab initio* pair creation, radiative processes, and nonideal effects. Including general relativity and radiative processes in this problem proved to be a necessary step, both to model the dynamics of the plasma and to reproduce observables.

In this thesis, I have presented global simulations of the intrinsic activity of spark gaps in the magnetosphere (Chap. 12). They have proved that the Blandford-Znajek mechanism can indeed be activated self-consistently in a magnetic monopole configuration. This reinforces the status of this process as a plausible energy extraction mechanism, a currently debated issue in the scientific community. I have also been able to pinpoint the location of the spark gaps and the separation surface of the electron/positron flow at the inner light surface, thereby answering a long-standing question in this field. These simulations have shown that pair creation occurs intermittently on small spatial scales. I have also performed simulations in a more realistic magnetic configuration, with paraboloidal field lines switching polarity at the equator (Chap. 13). These simulations display a reconnecting current sheet, allowing me to understand how the spark gap activity is affected. They have shown that magnetospheres can dissipate electromagnetic energy very efficiently through magnetic reconnection. Finally, I have coupled the kinetic code `Zeltron` to a public ray-tracing code, in order to produce synthetic gamma-ray lightcurves and millimeter images (Chap. 14). This has enabled me to bridge the gap between magnetospheric models and observations. I have found that although most of the dissipated energy is converted into observable inverse Compton radiation, no flare as intense as those observed in AGN can be reproduced. High-energy flares, if of magnetospheric origin, could therefore result from the response of the magnetosphere to external changes. I have also shown that the morphology of the EHT synchrotron images is consistent with a magnetospheric origin.

These studies still need improvements to yield conclusive answers. One issue in my gap simulations is the crude model of the ambient emission field. Not only does a realistic radiation field have a broad spectrum, but the isotropy hypothesis is also likely to be false, as FIDOs move at relativistic speeds and should observe beamed soft emission. It will be interesting to see whether a more realistic soft radiation field alters our findings. A bigger issue regarding these simulations is the limited separation of scales. Increasing it will be challenging, but it will allow me to ascertain my important conclusions, such as the role of the light surface. Performing a more extensive parameter study might help to understand the physics of pair production, depending on the optical depth and the spin of the black hole. In parallel, it might be instructive to downscale current 1D kinetic simulations of spark gaps in order to compare their results to my 2D radiative simulations. Indeed, there is currently tension regarding the roles of the null surface and inner light surface.

Concerning my paraboloid simulations, a major step forward will be to extend these models to 3D, in order to understand whether the cyclic dynamics of the current sheet can be sustained in a more realistic setup. Importantly, simulations with variable external parameters (the optical depth or the magnetic strength) should be investigated to see if they can reproduce flares. Rapid changes in the external conditions might result in stark rises in the high-energy luminosity, which is all the more likely as we have proved that the radiative efficiency of the current sheet is very high.

With recent EHT breakthroughs, such as the ability to perform polarization-sensitive measurements, and with the expectations of future improvements in the angular resolution of the observations, it will become even more important to predict finer features of the synchrotron images. Furthermore, the EHT collaboration has already recorded data from the supermassive black hole Sgr A*. This data has proved hard to analyze because of a large time variability and has not been released yet. Yet, dynamical imaging techniques are being developed to overcome this issue, so the importance of time-dependent predictions is growing. I intend to refine and develop the black-hole images produced by my magnetospheric simulations. I plan to study the time variability of these images and to include the transport of the Stokes parameters in the ray-tracing treatment, in order to produce polarized images. This will allow me to perform more decisive tests of the magnetospheric model. I also want to check the hypothesis that synchrotron cooling can successfully quench the polar cap hot spots, and more generally, whether synchrotron losses can affect the dynamics of the simulations.

So far, I have studied magnetic configurations in which the black hole and the accretion disk are disconnected, with a focus on the former. However, such configurations are disfavored by the recent GRAVITY observations of a hot spot orbiting Sgr A*, which are reminiscent of a solar flare. The black hole and the disk can be magnetically linked, in an analogous way to a young star and its protoplanetary disk. Magnetic reconnection between field lines connected to the black hole and open field lines threading the disk could produce a burst of energetic particles that would stream down the magnetic field lines and bombard the accretion disk, heating it and generating flares. In the future, I would like to reproduce these GRAVITY observations by studying a configuration with a magnetic linkage between the black hole and the accretion disk. The accretion flow can be simply modeled as a thin and perfectly conducting equatorial disk since it is not necessary to know the full dynamics of the flow in this problem. 2D axisymmetric solutions can be used to evaluate how much energy extracted from the black hole is transferred to the disk, depending on the spin of the black hole. To compare with GRAVITY observations, I will take advantage of my experience with ray tracing to reconstruct infra-red lightcurves. This specific problem will require 3D simulations.

To go even further in modeling the phenomenology of accreting black holes from first principles, I would also like to apply the techniques presented in this thesis to black-hole coronæ. Hard X-ray emission from X-ray binaries is thought to result from inverse Compton scattering of thermal soft X-rays by a hot and tenuous plasma. What mechanism heats this corona is still not completely understood, but magnetic reconnection is a serious candidate (Beloborodov, 2017). In this environment, reconnection could be driven by the dynamics of twisted magnetic loops with footpoints anchored in the disk (Uzdensky & Goodman, 2008). Configurations with twisted magnetic field lines are ubiquitous in high-energy astrophysical systems. Theoretical and numerical studies in the MHD framework have shown

that when a flux tube is twisted so that the toroidal and poloidal components of its magnetic field become comparable, it becomes unstable and gets disrupted. This twist can occur spontaneously (e.g., through the kink instability), but in this context, it can be triggered by the relative motions of the footpoints. As the field lines break up and reconnect, magnetic reconnection is triggered, dissipating magnetic energy and ejecting hot plasma and energetic particles. In a previous study of interacting pulsar magnetospheres, I was already confronted withs twisted magnetic fields (Crinquand et al., 2019), but I was unable to fully capture the consequences of that twist at the time, as I was performing axisymmetric simulations. In the longer term, my goal will be to perform 3D simulations of a black-hole corona, in order to test this magnetic reconnection scenario.

The main shortcoming of general relativistic global kinetic simulations is their relatively poor separation of scales, which is limited by the need to resolve plasma kinetic scales. This fact makes it clear that future breakthroughs in numerical modeling of black-hole magnetospheres will involve a coupling between kinetic and fluid/force-free approaches, which are cruder but computationally much cheaper. That will allow to scale up simulations to more realistic dynamic ranges. Such an approach was initiated by Makwana et al. (2017) but has not been applied to global simulations yet. Ultimately, this might establish a connection between numerical simulations of accretion flows (electron/proton plasma, large scales) and the inner magnetosphere (pair plasma, inner scales). Achieving the ambitious goals previously outlined will necessitate a future increase in computer power. With the performances of single-core CPUs currently reaching a plateau, this will probably require porting astrophysical codes to GPU-based architectures.

From the observational point of view, there are several milestones ahead of us. In the VHE band, the future CTA observatory will be able to detect much fainter sources with improved angular resolution. This might allow us to resolve the VHE source in the Galactic center and to broaden our sample of VHE-loud radio galaxies. Besides, the initial achievements of the EHT collaboration are very promising for the prospects of high-resolution radio interferometry. When analyzed and released, observations from Sgr A* will offer a glimpse of the time variability of black-hole magnetospheres. Besides, as more radio telescopes will join the array and with improved frequency and resolution, it will become possible to image both the inner magnetosphere and the launching region of the jet on the same image. Hopefully, this will allow us to grasp the long-sought connection between the black hole and the jet, for instance by comparing the outgoing direction of the jet at the scale of the magnetosphere with the large-scale jet. In retrospect, I can only stand by what I said in the opening foreword: right now, it is extremely exciting to work in the field of relativistic astrophysics.

Bibliography

Abbott, B. P., Abbott, R., Abbott, T. D., et al., *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett., 116, 061102 (2016). (Cited on page 10.)

Abbott, B. P., Abbott, R., Abbott, T. D., et al., *The basic physics of the binary black hole merger GW150914*, Annalen der Physik, 529, 1600209 (2017). (Cited on page 10.)

Acciari, V. A., Aliu, E., Arlen, T., et al., *Radio Imaging of the Very-High-Energy γ-Ray Emission Region in the Central Engine of a Radio Galaxy*, Science, 325, 444 (2009). (Cited on pages 35 and 36.)

Acero, F., Aharonian, F., Akhperjanian, A. G., et al., *Localizing the VHE γ-ray source at the Galactic Centre*, MNRAS, 1877 (2010). (Cited on page 35.)

Aharonian, F., Akhperjanian, A. G., Anton, G., et al., *Discovery of Very High Energy γ-Ray Emission from Centaurus a with H.E.S.S.*, ApJ, 695, L40 (2009). (Cited on page 34.)

Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al., *An Exceptional Very High Energy Gamma-Ray Flare of PKS 2155-304*, ApJ, 664, L71 (2007). (Cited on pages 33, 34, and 196.)

Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al., *Fast Variability of Tera-Electron Volt γ Rays from the Radio Galaxy M87*, Science, 314, 1424 (2006). (Cited on pages 34 and 35.)

Aharonian, F. A., Atoian, A. M., & Nagapetian, A. M., *Photoproduction of electron-positron pairs in compact X-ray sources*, Astrofizika, 19, 323 (1983). (Cited on page 112.)

Aharonian, F. A., Barkov, M. V., & Khangulyan, D., *Scenarios for Ultrafast Gamma-Ray Variability in AGN*, ApJ, 841, 61 (2017). (Cited on page 37.)

Albert, J., Aliu, E., Anderhub, H., et al., *Variable Very High Energy γ-Ray Emission from Markarian 501*, ApJ, 669, 862 (2007). (Cited on page 33.)

Aleksić, J., Alvarez, E. A., Antonelli, L. A., et al., *Detection of very-high energy γ-ray emission from NGC 1275 by the MAGIC telescopes*, A&A, 539, L2 (2012). (Cited on page 34.)

Aleksić, J., Ansoldi, S., Antonelli, L. A., et al., *Black hole lightning due to particle acceleration at subhorizon scales*, Science, 346, 1080 (2014). (Cited on page 34.)

Algaba, J. C., Anczarski, J., Asada, K., et al., *Broadband Multi-wavelength Properties of M87 during the 2017 Event Horizon Telescope Campaign*, arXiv e-prints, arXiv:2104.06855 (2021). (Cited on pages 26 and 27.)

Aliu, E., Arlen, T., Aune, T., et al., *VERITAS Observations of Day-scale Flaring of M 87 in 2010 April*, ApJ, 746, 141 (2012). (Cited on pages 34 and 35.)

Amaro-Seoane, P., Audley, H., Babak, S., et al., *Laser Interferometer Space Antenna*, arXiv e-prints, arXiv:1702.00786 (2017). (Cited on page 17.)

Archer, A., Barnacka, A., Beilicke, M., et al., *Very-high Energy Observations of the Galactic Center Region by VERITAS in 2010-2012*, ApJ, 790, 149 (2014). (Cited on page 35.)

Archer, A., Benbow, W., Bird, R., et al., *VERITAS Discovery of VHE Emission from the Radio Galaxy 3C 264: A Multiwavelength Study*, ApJ, 896, 41 (2020). (Cited on page 34.)

Asada, K., Nakamura, M., Doi, A., Nagai, H., & Inoue, M., *Discovery of Sub- to Superluminal Motions in the M87 Jet: An Implication of Acceleration from Sub-relativistic to Relativistic Speeds*, ApJ, 781, L2 (2014). (Cited on page 21.)

Baade, W. & Minkowski, R., *Identification of the Radio Sources in Cassiopeia, Cygnus A, and Puppis A.*, ApJ, 119, 206 (1954). (Cited on page 11.)

Baade, W. & Zwicky, F., *On Super-novae*, Proceedings of the National Academy of Science, 20, 254 (1934). (Cited on page 4.)

Bacchini, F., Ripperda, B., Chen, A. Y., & Sironi, L., *Generalized, Energy-conserving Numerical Simulations of Particles in General Relativity. I. Time-like and Null Geodesics*, ApJS, 237, 6 (2018). (Cited on page 134.)

Bacchini, F., Ripperda, B., Porth, O., & Sironi, L., *Generalized, Energy-conserving Numerical Simulations of Particles in General Relativity. II. Test Particles in Electromagnetic Fields and GRMHD*, ApJS, 240, 40 (2019). (Cited on page 134.)

Baganoff, F. K., Maeda, Y., Morris, M., et al., *Chandra X-Ray Spectroscopic Imaging of Sagittarius A* and the Central Parsec of the Galaxy*, ApJ, 591, 891 (2003). (Cited on pages 30 and 31.)

Bai, X.-N., Caprioli, D., Sironi, L., & Spitkovsky, A., *Magnetohydrodynamic-particle-in-cell Method for Coupling Cosmic Rays with a Thermal Plasma: Application to Non-relativistic Shocks*, ApJ, 809, 55 (2015). (Cited on page 122.)

Bardeen, J. M. & Petterson, J. A., *The Lense-Thirring Effect and Accretion Disks around Kerr Black Holes*, ApJ, 195, L65 (1975). (Cited on page 14.)

Bardeen, J. M., Press, W. H., & Teukolsky, S. A., *Rotating Black Holes: Locally Nonrotating Frames, Energy Extraction, and Scalar Synchrotron Radiation*, ApJ, 178, 347 (1972). (Cited on pages 186 and 205.)

Barret, D., McClintock, J. E., & Grindlay, J. E., *Luminosity Differences between Black Holes and Neutron Stars*, ApJ, 473, 963 (1996). (Cited on page 9.)

Barrière, N. M., Tomsick, J. A., Baganoff, F. K., et al., *NuSTAR Detection of High-energy X-Ray Emission and Rapid Variability from Sagittarius A* Flares*, ApJ, 786, 46 (2014). (Cited on page 32.)

Begelman, M. & Rees, M., *Gravity's Fatal Attraction: Black Holes in the Universe*, 2nd edn., Cambridge University Press (2009). (Cited on pages 3, 8, and 15.)

Begelman, M. C., Fabian, A. C., & Rees, M. J., *Implications of very rapid TeV variability in blazars*, MNRAS, 384, L19 (2008). (Cited on page 33.)

Bekefi, G., *Radiation processes in plasmas*, 1st edn., John Wiley & Sons (1966). (Cited on pages 115 and 116.)

Beloborodov, A. M., *Radiative Magnetic Reconnection Near Accreting Black Holes*, ApJ, 850, 141 (2017). (Cited on pages 18 and 212.)

Belyaev, M. A., *PICsar: A 2.5D axisymmetric, relativistic, electromagnetic, Particle in Cell code with a radiation absorbing boundary*, New A, 36, 37 (2015a). (Cited on page 139.)

Belyaev, M. A., *Dissipation, energy transfer, and spin-down luminosity in 2.5D PIC simulations of the pulsar magnetosphere*, MNRAS, 449, 2759 (2015b). (Cited on page 44.)

Bertin, M., Faroux, J.-P., & Renault, J., *Électromagnétisme 4 : milieux diélectriques et milieux aimantés*, 1st edn., Dunod (1996). (Cited on page 68.)

Bertschinger, E. (1999), *Hamiltonian Dynamics of Particle Motion*, URL: `http://web.mit.edu/edbert/GR/gr3.pdf`. Last visited on 2021/03/10. (Cited on page 145.)

Beskin, V. S., *Magnetohydrodynamic models of astrophysical jets*, Physics Uspekhi, 53, 1199 (2010). (Cited on page 95.)

Beskin, V. S., *MHD Flows in Compact Astrophysical Objects: Accretion, Winds and Jets*, 1st edn., Springer (2010). (Cited on pages 91 and 93.)

Beskin, V. S., Istomin, Y. N., & Parev, V. I., *Filling the Magnetosphere of a Supermassive Black-Hole with Plasma*, Soviet Ast., 36, 642 (1992). (Cited on page 39.)

Bicak, J. & Janis, V., *Magnetic fluxes across black holes*, MNRAS, 212, 899 (1985). (Cited on pages 74, 80, and 82.)

Bilous, A. V., Watts, A. L., Harding, A. K., et al., *A NICER View of PSR J0030+0451: Evidence for a Global-scale Multipolar Magnetic Field*, ApJ, 887, L23 (2019). (Cited on page 89.)

Birdsall, C. K. & Langdon, A. B., *Plasma Physics via Computer Simulation*, 1st edn., McGraw-Hill (1985). (Cited on pages 121, 122, 125, 126, 130, and 131.)

Bičák, J., Karas, V., & Ledvinka, T. (2007), in *Black Holes from Stars to Galaxies – Across the Range of Masses*, ed., V. Karas & G. Matt, Vol. 238, Cambridge Univeristy Press, 139–144. (Cited on page 81.)

Blandford, R., Meier, D., & Readhead, A., *Relativistic Jets from Active Galactic Nuclei*, ARA&A, 57, 467 (2019). (Cited on pages 18 and 27.)

Blandford, R. D. & Begelman, M. C., *On the fate of gas accreting at a low rate on to a black hole*, MNRAS, 303, L1 (1999). (Cited on page 24.)

Blandford, R. D. & Königl, A., *Relativistic jets as compact radio sources.*, ApJ, 232, 34 (1979). (Cited on pages 25 and 27.)

Blandford, R. D. & McKee, C. F., *Reverberation mapping of the emission line regions of Seyfert galaxies and quasars.*, ApJ, 255, 419 (1982). (Cited on page 16.)

Blandford, R. D. & Payne, D. G., *Hydromagnetic flows from accretion disks and the production of radio jets.*, MNRAS, 199, 883 (1982). (Cited on page 22.)

Blandford, R. D. & Znajek, R. L., *Electromagnetic extraction of energy from Kerr black holes*, MNRAS, 179, 433 (1977). (Cited on pages 23, 37, 38, 41, 84, 91, 94, and 101.)

Blau, M. (2018), *Lecture Notes on General Relativity*, URL: `http://www.blau.itp.unibe.ch/newlecturesGR.pdf`. Last visited on 2021/03/05. (Cited on pages 49, 65, 71, 75, 137, and 143.)

Blumenthal, G. R. & Gould, R. J., *Bremsstrahlung, Synchrotron Radiation, and Compton Scattering of High-Energy Electrons Traversing Dilute Gases*, Reviews of Modern Physics, 42, 237 (1970). (Cited on pages 105, 107, 113, 114, and 118.)

Bolton, C. T., *Identification of Cygnus X-1 with HDE 226868*, Nature, 235, 271 (1972). (Cited on page 9.)

Bondi, H., *On spherically symmetrical accretion*, MNRAS, 112, 195 (1952). (Cited on page 24.)

Bonometto, S. & Rees, M. J., *On possible observable effects of electron pair-production in QSOs*, MNRAS, 152, 21 (1971). (Cited on page 111.)

Boris, J. P. (1970), in *Proceedings of the Fourth Conference on Numerical Simulations of Plasmas*, ed., J. P. Boris, Naval Research Laboratory, 3–68. (Cited on page 125.)

Brodatzki, K. A., Pardy, D. J. S., Becker, J. K., & Schlickeiser, R., *Internal $\gamma\gamma$ Opacity in Active Galactic Nuclei and the Consequences for the TeV Observations of M87 and Cen A*, ApJ, 736, 98 (2011). (Cited on page 39.)

Broderick, A. E., Loeb, A., & Narayan, R., *The Event Horizon of Sagittarius A**, ApJ, 701, 1357 (2009). (Cited on page 31.)

Broderick, A. E. & Tchekhovskoy, A., *Horizon-scale Lepton Acceleration in Jets: Explaining the Compact Radio Emission in M87*, ApJ, 809, 97 (2015). (Cited on page 40.)

Bronzwaer, T., Davelaar, J., Younsi, Z., et al., *Visibility of black hole shadows in low-luminosity AGN*, MNRAS, 501, 4722 (2021). (Cited on page 199.)

Burbidge, G. R., *On Synchrotron Radiation from Messier 87.*, ApJ, 124, 416 (1956). (Cited on page 11.)

Burbidge, G. R., *Estimates of the Total Energy in Particles and Magnetic Field in the Non-Thermal Radio Sources.*, ApJ, 129, 849 (1959). (Cited on page 11.)

Carilli, C. L. & Barthel, P. D., *Cygnus A*, A&A Rev., 7, 1 (1996). (Cited on page 12.)

Carroll, S. M., *Spacetime and Geometry: An Introduction to General Relativity*, Pearson (2003). (Cited on page 49.)

Carter, B., *Global Structure of the Kerr Family of Gravitational Fields*, Physical Review, 174, 1559 (1968). (Cited on page 186.)

Carter, B., *Axisymmetric Black Hole Has Only Two Degrees of Freedom*, Phys. Rev. Lett., 26, 331 (1971). (Cited on page 51.)

Celotti, A. & Fabian, A. C., *The kinetic power and luminosity of parsec-scale radio jets - an argument for heavy jets.*, MNRAS, 264, 228 (1993). (Cited on page 23.)

Cerutti, B. (2010), High-energy gamma-ray emission in compact binaries, Theses, Université de Grenoble, URL: `https://tel.archives-ouvertes.fr/tel-00499281`. (Cited on page 110.)

Cerutti, B. & Beloborodov, A. M., *Electrodynamics of Pulsar Magnetospheres*, Space Sci. Rev., 207, 111 (2017). (Cited on page 91.)

Cerutti, B., Philippov, A., Parfrey, K., & Spitkovsky, A., *Particle acceleration in axisymmetric pulsar current sheets*, MNRAS, 448, 606 (2015). (Cited on pages 138, 139, 143, and 159.)

Cerutti, B., Philippov, A. A., & Dubus, G., *Dissipation of the striped pulsar wind and non-thermal particle acceleration: 3D PIC simulations*, A&A, 642, A204 (2020). (Cited on page 21.)

Cerutti, B., Philippov, A. A., & Spitkovsky, A., *Modelling high-energy pulsar light curves from first principles*, MNRAS, 457, 2401 (2016). (Cited on pages 114, 118, 138, 139, 185, and 191.)

Cerutti, B., Uzdensky, D. A., & Begelman, M. C., *Extreme Particle Acceleration in Magnetic Reconnection Layers: Application to the Gamma-Ray Flares in the Crab Nebula*, ApJ, 746, 148 (2012). (Cited on pages 118 and 193.)

Cerutti, B., Werner, G. R., Uzdensky, D. A., & Begelman, M. C., *Simulations of Particle Acceleration beyond the Classical Synchrotron Burnoff Limit in Magnetic Reconnection: An Explanation of the Crab Flares*, ApJ, 770, 147 (2013). (Cited on page 181.)

Chan, C.-k., Psaltis, D., & Özel, F., *GRay: A Massively Parallel GPU-based Code for Ray Tracing in Relativistic Spacetimes*, ApJ, 777, 13 (2013). (Cited on page 189.)

Chandrasekhar, S., *The Maximum Mass of Ideal White Dwarfs*, ApJ, 74, 81 (1931). (Cited on page 4.)

Chandrasekhar, S., *The highly collapsed configurations of a stellar mass (Second paper)*, MNRAS, 95, 207 (1935). (Cited on page 4.)

Chen, A. Y. & Yuan, Y., *Physics of Pair Producing Gaps in Black Hole Magnetospheres. II. General Relativity*, ApJ, 895, 121 (2020). (Cited on pages 43, 167, and 171.)

Chen, A. Y., Yuan, Y., & Yang, H., *Physics of Pair Producing Gaps in Black Hole Magnetospheres*, ApJ, 863, L31 (2018). (Cited on page 43.)

Cheng, K. S., Ruderman, M., & Zhang, L., *A Three-dimensional Outer Magnetospheric Gap Model for Gamma-Ray Pulsars: Geometry, Pair Production, Emission Morphologies, and Phase-resolved Spectra*, ApJ, 537, 964 (2000). (Cited on page 39.)

Christie, I. M., Petropoulou, M., Sironi, L., & Giannios, D., *Radiative signatures of plasmoid-dominated reconnection in blazar jets*, MNRAS, 482, 65 (2019). (Cited on page 193.)

Chruściel, P. T., Costa, J. L., & Heusler, M., *Stationary Black Holes: Uniqueness and Beyond*, Living Reviews in Relativity, 15, 7 (2012). (Cited on page 51.)

Cohen, M. H., Lister, M. L., Homan, D. C., et al., *Relativistic Beaming and the Intrinsic Properties of Extragalactic Radio Jets*, ApJ, 658, 232 (2007). (Cited on page 19.)

Colgate, S. A. & White, R. H., *The Hydrodynamic Behavior of Supernovae Explosions*, ApJ, 143, 626 (1966). (Cited on page 5.)

Collectif Jolidon, *Physique expérimentale: Optique, magnétisme, électrotechnique, mécanique, thermodynamique et physique non linéaire*, 1st edn., EDP Sciences (2021). (Cited on page 81.)

Conticini, P., Boé, P., & Bloch-Lainé, J.-L., *Sensations. 288 recettes de pâtisseries*, 1st edn., La Martinière (2012). (Cited on page 240.)

Coroniti, F. V., *Magnetically Striped Relativistic Magnetohydrodynamic Winds: The Crab Nebula Revisited*, ApJ, 349, 538 (1990). (Cited on page 182.)

Corral-Santana, J. M., Casares, J., Muñoz-Darias, T., et al., *BlackCAT: A catalogue of stellar-mass black holes in X-ray transients*, A&A, 587, A61 (2016). (Cited on page 10.)

Crinquand, B., Cerutti, B., Philippov, A., Parfrey, K., & Dubus, G., *Multidimensional Simulations of Ergospheric Pair Discharges around Black Holes*, Phys. Rev. Lett., 124, 145101 (2020). (Cited on pages 45, 149, and 159.)

Crinquand, B., Cerutti, B., & Dubus, G., *Kinetic modeling of the electromagnetic precursor from an axisymmetric binary pulsar coalescence*, A&A, 622, A161 (2019). (Cited on page 213.)

Crinquand, B., Cerutti, B., Dubus, G., Parfrey, K., & Philippov, A., *Synthetic gamma-ray light curves of Kerr black hole magnetospheric activity from particle-in-cell simulations*, A&A, 650, A163 (2021). (Cited on pages 45, 172, and 185.)

Cruz, F., Grismayer, T., & Silva, L. O., *Kinetic Model of Large-amplitude Oscillations in Neutron Star Pair Cascades*, ApJ, 908, 149 (2021). (Cited on page 43.)

Cui, W., Heindl, W. A., Rothschild, R. E., et al., *Rossi X-Ray Timing Explorer Observation of Cygnus X-1 in Its High State*, ApJ, 474, L57 (1997). (Cited on page 8.)

Cunningham, C. T. & Bardeen, J. M., *The Optical Appearance of a Star Orbiting an Extreme Kerr Black Hole*, ApJ, 183, 237 (1973). (Cited on pages 187, 188, and 189.)

De Villiers, J.-P., Hawley, J. F., & Krolik, J. H., *Magnetically Driven Accretion Flows in the Kerr Metric. I. Models and Overall Structure*, ApJ, 599, 1238 (2003). (Cited on page 42.)

Derouillat, J., Beck, A., Pérez, F., et al., *SMILEI : A collaborative, open-source, multi-purpose particle-in-cell code for plasma simulation*, Computer Physics Communications, 222, 351 (2018). (Cited on page 133.)

Dexter, J. & Agol, E., *A Fast New Public Code for Computing Photon Orbits in a Kerr Spacetime*, ApJ, 696, 1616 (2009). (Cited on pages 185, 189, and 190.)

Dexter, J., Tchekhovskoy, A., Jiménez-Rosales, A., et al., *Sgr A* near-infrared flares from reconnection events in a magnetically arrested disc*, MNRAS, 497, 4999 (2020). (Cited on page 32.)

Di Matteo, T., Allen, S. W., Fabian, A. C., Wilson, A. S., & Young, A. J., *Accretion onto the Supermassive Black Hole in M87*, ApJ, 582, 133 (2003). (Cited on page 24.)

Dirac, P. A. M., *On the Annihilation of Electrons and Protons*, Proceedings of the Cambridge Philosophical Society, 26, 361 (1930). (Cited on page 113.)

Dodds-Eden, K., Porquet, D., Trap, G., et al., *Evidence for X-Ray Synchrotron Emission from Simultaneous Mid-Infrared to X-Ray Observations of a Strong Sgr A* Flare*, ApJ, 698, 676 (2009). (Cited on pages 31 and 32.)

Dodin, I. Y. & Fisch, N. J., *Vlasov equation and collisionless hydrodynamics adapted to curved spacetime*, Physics of Plasmas, 17, 112118 (2010). (Cited on pages 60 and 134.)

Dodson, R., Edwards, P. G., & Hirabayashi, H., *Milliarcsecond-Scale Spectral Properties and Jet Motions in M 87*, PASJ, 58, 243 (2006). (Cited on page 35.)

Doeleman, S. S., Weintroub, J., Rogers, A. E. E., et al., *Event-horizon-scale structure in the supermassive black hole candidate at the Galactic Centre*, Nature, 455, 78 (2008). (Cited on page 31.)

Drenkhahn, G. & Spruit, H. C., *Efficient acceleration and radiation in Poynting flux powered GRB outflows*, A&A, 391, 1141 (2002). (Cited on page 182.)

Eckart, A., Baganoff, F. K., Schödel, R., et al., *The flare activity of Sagittarius A*. New coordinated mm to X-ray observations*, A&A, 450, 535 (2006). (Cited on page 31.)

Einstein, A., *On a Stationary System with Spherical Symmetry Consisting of Many Gravitating Masses*, Annals of Mathematics, 40, 922 (1939). (Cited on page 3.)

Elvis, M., Wilkes, B. J., McDowell, J. C., et al., *Atlas of Quasar Energy Distributions*, ApJS, 95, 1 (1994). (Cited on page 18.)

Endean, V. G., *Lorentz Force-Free Pulsar Rotating Fields*, ApJ, 187, 359 (1974). (Cited on page 90.)

Esirkepov, T. Z., *Exact charge conservation scheme for Particle-in-Cell simulation with an arbitrary form-factor*, Computer Physics Communications, 135, 144 (2001). (Cited on page 131.)

Evans, D. A., Kraft, R. P., Worrall, D. M., et al., *Chandra and XMM-Newton Observations of the Nucleus of Centaurus A*, ApJ, 612, 786 (2004). (Cited on page 35.)

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al., *First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole*, ApJ, 875, L1 (2019a). (Cited on pages 28 and 29.)

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al., *First M87 Event Horizon Telescope Results. V. Physical Origin of the Asymmetric Ring*, ApJ, 875, L5 (2019b). (Cited on pages 28, 42, 94, 161, and 185.)

Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al., *First M87 Event Horizon Telescope Results. VI. The Shadow and Mass of the Central Black Hole*, ApJ, 875, L6 (2019c). (Cited on page 28.)

Event Horizon Telescope Collaboration, Akiyama, K., Algaba, J. C., et al., *First M87 Event Horizon Telescope Results. VII. Polarization of the Ring*, ApJ, 910, L12 (2021a). (Cited on pages 28 and 29.)

Event Horizon Telescope Collaboration, Akiyama, K., Algaba, J. C., et al., *First M87 Event Horizon Telescope Results. VIII. Magnetic Field Structure near The Event Horizon*, ApJ, 910, L13 (2021b). (Cited on pages 28 and 185.)

Fabian, A. C. & Canizares, C. R., *Do massive black holes reside in elliptical galaxies?*, Nature, 333, 829 (1988). (Cited on page 24.)

Fabian, A. C., Iwasawa, K., Reynolds, C. S., & Young, A. J., *Broad Iron Lines in Active Galactic Nuclei*, PASP, 112, 1145 (2000). (Cited on page 16.)

Falcke, H., Melia, F., & Agol, E., *Viewing the Shadow of the Black Hole at the Galactic Center*, ApJ, 528, L13 (2000). (Cited on pages 28 and 198.)

Fernandes, C. A. C., Jarvis, M. J., Rawlings, S., et al., *Evidence for a maximum jet efficiency for the most powerful radio galaxies*, MNRAS, 411, 1909 (2011). (Cited on pages 19 and 22.)

Feynman, R. P., Leighton, R. B., & Sands, M. L., *The Feynman lectures on physics, volume 1: Mainly mechanics, radiation, and heat*, 1st edn., Addison-Wesley (1963). (Cited on page 123.)

Finkelstein, D., *Past-Future Asymmetry of the Gravitational Field of a Point Particle*, Physical Review, 110, 965 (1958). (Cited on page 5.)

Fitch, W. S., Pacholczyk, A. G., & Weymann, R. J., *Light Variations of the Seyfert Galaxy NGC 4151*, ApJ, 150, L67 (1967). (Cited on page 12.)

Fitzpatrick, R., *Plasma Physics: An Introduction*, 1st edn., CRC Press (2015). (Cited on pages 42 and 122.)

Font, J. A., Ibáñez, J. M., & Papadopoulos, P., *Non-axisymmetric relativistic Bondi-Hoyle accretion on to a Kerr black hole*, MNRAS, 305, 920 (1999). (Cited on page 57.)

Ford, H. C., Harms, R. J., Tsvetanov, Z. I., et al., *Narrowband HST Images of M87: Evidence for a Disk of Ionized Gas around a Massive Black Hole*, ApJ, 435, L27 (1994). (Cited on page 16.)

Forest, E. & Ruth, R. D., *Fourth-order symplectic integration*, Physica D Nonlinear Phenomena, 43, 105 (1990). (Cited on page 125.)

Foucart, F., Chandra, M., Gammie, C. F., & Quataert, E., *Evolution of accretion discs around a Kerr black hole using extended magnetohydrodynamics*, MNRAS, 456, 1332 (2016). (Cited on page 43.)

Frolov, V. P. & Novikov, I. D., *Black Hole Physics: Basic Concepts and New Developments*, 1st edn., Springer (1998). (Cited on pages 51 and 79.)

Frolov, V. P. & Zelnikov, A., *Introduction to Black Hole Physics*, 1st edn., Oxford University Press (2011). (Cited on pages 49, 52, 187, 190, 205, and 207.)

Gammie, C. F., McKinney, J. C., & Tóth, G., *HARM: A Numerical Scheme for General Relativistic Magnetohydrodynamics*, ApJ, 589, 444 (2003). (Cited on page 41.)

Garing, C., *Magnétisme : statique, induction et milieux*, 1st edn., Ellipses (1999). (Cited on page 95.)

Garrington, S. T., Leahy, J. P., Conway, R. G., & Laing, R. A., *A systematic asymmetry in the polarization properties of double radio sources with one jet*, Nature, 331, 147 (1988). (Cited on page 19.)

Gebhardt, K., Adams, J., Richstone, D., et al., *The Black Hole Mass in M87 from Gemini/NIFS Adaptive Optics Observations*, ApJ, 729, 119 (2011). (Cited on page 28.)

Genzel, R., Eisenhauer, F., & Gillessen, S., *The Galactic Center massive black hole and nuclear star cluster*, Reviews of Modern Physics, 82, 3121 (2010). (Cited on pages 17 and 30.)

Genzel, R., Thatte, N., Krabbe, A., Kroker, H., & Tacconi-Garman, L. E., *The Dark Mass Concentration in the Central Parsec of the Milky Way*, ApJ, 472, 153 (1996). (Cited on page 29.)

Ghez, A. M., Morris, M., Becklin, E. E., Tanner, A., & Kremenek, T., *The accelerations of stars orbiting the Milky Way's central black hole*, Nature, 407, 349 (2000). (Cited on page 29.)

Ghisellini, G., Tavecchio, F., Foschini, L., et al., *General physical properties of bright Fermi blazars*, MNRAS, 402, 497 (2010). (Cited on page 22.)

Ghisellini, G., Tavecchio, F., Maraschi, L., Celotti, A., & Sbarrato, T., *The power of relativistic jets is larger than the luminosity of their accretion disks*, Nature, 515, 376 (2014). (Cited on page 22.)

Giacconi, R., Gursky, H., Paolini, F. R., & Rossi, B. B., *Evidence for x Rays From Sources Outside the Solar System*, Phys. Rev. Lett., 9, 439 (1962). (Cited on page 8.)

Giannios, D., Uzdensky, D. A., & Begelman, M. C., *Fast TeV variability in blazars: jets in a jet*, MNRAS, 395, L29 (2009). (Cited on page 37.)

Giannios, D., Uzdensky, D. A., & Begelman, M. C., *Fast TeV variability from misaligned minijets in the jet of M87*, MNRAS, 402, 1649 (2010). (Cited on page 37.)

Gillessen, S., Eisenhauer, F., Trippe, S., et al., *Monitoring Stellar Orbits Around the Massive Black Hole in the Galactic Center*, ApJ, 692, 1075 (2009). (Cited on pages 29 and 30.)

Ginzburg, V. L. (1982), in *Classics in Radio Astronomy*, Vol. 10, Springer, 93. (Cited on page 11.)

Ginzburg, V. L. & Syrovatskii, S. I., *Cosmic Magnetobremsstrahlung (synchrotron Radiation)*, ARA&A, 3, 297 (1965). (Cited on page 113.)

Globus, N. & Levinson, A., *Loaded magnetohydrodynamic flows in Kerr spacetime*, Phys. Rev. D, 88, 084046 (2013). (Cited on pages 40 and 97.)

Goddi, C., Martí-Vidal, I., Messias, H., et al., *Polarimetric Properties of Event Horizon Telescope Targets from ALMA*, ApJ, 910, L14 (2021). (Cited on page 30.)

Goldreich, P. & Julian, W. H., *Pulsar Electrodynamics*, ApJ, 157, 869 (1969). (Cited on page 86.)

Goldstein, H., Poole Jr., C., & Safko, J., *Classical mechanics*, 3rd edn., Pearson (2012). (Cited on page 124.)

Goodman, J. & Uzdensky, D., *Reconnection in Marginally Collisionless Accretion Disk Coronae*, ApJ, 688, 555 (2008). (Cited on page 182.)

Gould, R. J. & Schréder, G. P., *Pair Production in Photon-Photon Collisions*, Physical Review, 155, 1404 (1967). (Cited on page 110.)

Gourgoulhon, E., *Relativité restreinte : Des Particules à l'Astrophysique*, CNRS Éditions (2010). (Cited on pages 56, 65, 71, and 116.)

Gourgoulhon, E., *3+1 Formalism in General Relativity*, Springer (2012). (Cited on pages 60, 62, 63, 77, and 80.)

Gourgoulhon, E. (2021), *Geometry and physics of black holes*, URL: `https://luth.obspm.fr/~luthier/gourgoulhon/bh16/bholes.pdf`. Last visited on 2021/03/05. (Cited on pages 49 and 57.)

Gralla, S. E., Holz, D. E., & Wald, R. M., *Black hole shadows, photon rings, and lensing rings*, Phys. Rev. D, 100, 024018 (2019). (Cited on page 198.)

Gralla, S. E., Lupsasca, A., & Strominger, A., *Observational signature of high spin at the Event Horizon Telescope*, MNRAS, 475, 3829 (2018). (Cited on pages 187 and 189.)

Gravity Collaboration, Abuter, R., Amorim, A., et al., *Detection of orbital motions near the last stable circular orbit of the massive black hole SgrA**, A&A, 618, L10 (2018a). (Cited on page 32.)

Gravity Collaboration, Abuter, R., Amorim, A., et al., *Detection of the gravitational redshift in the orbit of the star S2 near the Galactic centre massive black hole*, A&A, 615, L15 (2018b). (Cited on pages 17 and 30.)

Gravity Collaboration, Abuter, R., Amorim, A., et al., *Detection of the Schwarzschild precession in the orbit of the star S2 near the Galactic centre massive black hole*, A&A, 636, L5 (2020a). (Cited on pages 17 and 30.)

Gravity Collaboration, Abuter, R., Amorim, A., et al., *The flux distribution of Sgr A**, A&A, 638, A2 (2020b). (Cited on page 31.)

Gravity Collaboration, Bauböck, M., Dexter, J., et al., *Modeling the orbital motion of Sgr A*'s near-infrared flares*, A&A, 635, A143 (2020c). (Cited on page 32.)

Gravity Collaboration, Jiménez-Rosales, A., Dexter, J., et al., *Dynamically important magnetic fields near the event horizon of Sgr A**, A&A, 643, A56 (2020d). (Cited on page 32.)

Gültekin, K., Richstone, D. O., Gebhardt, K., et al., *The M-σ and M-L Relations in Galactic Bulges, and Determinations of Their Intrinsic Scatter*, ApJ, 698, 198 (2009). (Cited on page 15.)

Gunn, J. E., *On the Distances of the Quasi-Stellar Objects*, ApJ, 164, L113 (1971). (Cited on page 13.)

H. E. S. S. Collaboration, Abdalla, H., Abramowski, A., et al., *Characterizing the γ-ray long-term variability of PKS 2155-304 with H.E.S.S. and Fermi-LAT*, A&A, 598, A39 (2017). (Cited on page 196.)

Haardt, F. & Maraschi, L., *X-Ray Spectra from Two-Phase Accretion Disks*, ApJ, 413, 507 (1993). (Cited on page 18.)

Hada, K., Doi, A., Kino, M., et al., *An origin of the radio jet in M87 at the location of the central black hole*, Nature, 477, 185 (2011). (Cited on pages 27 and 28.)

Hairer, E., Lubich, C., & Wanner, G., *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd edn., Springer (2005). (Cited on pages 125 and 136.)

Haller, J. W., Rieke, M. J., Rieke, G. H., et al., *Stellar Kinematics and the Black Hole in the Galactic Center*, ApJ, 456, 194 (1996). (Cited on page 29.)

Harding, A. K. & Lai, D., *Physics of strongly magnetized neutron stars*, Reports on Progress in Physics, 69, 2631 (2006). (Cited on page 89.)

Hasinger, G., Cappelluti, N., Brunner, H., et al., *The XMM-Newton Wide-Field Survey in the COSMOS Field. I. Survey Description*, ApJS, 172, 29 (2007). (Cited on page 9.)

Hazard, C., Mackey, M. B., & Shimmins, A. J., *Investigation of the Radio Source 3C 273 By The Method of Lunar Occultations*, Nature, 197, 1037 (1963). (Cited on page 12.)

Henri, G. & Pelletier, G., *Relativistic Electron-Positron Beam Formation in the Framework of the Two-Flow Model for Active Galactic Nuclei*, ApJ, 383, L7 (1991). (Cited on page 23.)

Henri, G. & Saugé, L., *The Bulk Lorentz Factor Crisis of TeV Blazars: Evidence for an Inhomogeneous Pileup Energy Distribution?*, ApJ, 640, 185 (2006). (Cited on page 33.)

HESS Collaboration, Abdalla, H., Abramowski, A., et al., *H.E.S.S. discovery of very high energy γ-ray emission from PKS 0625-354*, MNRAS, 476, 4187 (2018). (Cited on page 34.)

Hirotani, K., Iguchi, S., Kimura, M., & Wajima, K., *Pair Plasma Dominance in the 3C 279 Jet on Parsec Scales*, PASJ, 51, 263 (1999). (Cited on page 23.)

Hirotani, K. & Okamoto, I., *Pair Plasma Production in a Force-free Magnetosphere around a Supermassive Black Hole*, ApJ, 497, 563 (1998). (Cited on page 39.)

Hirotani, K. & Pu, H.-Y., *Energetic Gamma Radiation from Rapidly Rotating Black Holes*, ApJ, 818, 50 (2016). (Cited on pages 37, 38, 39, and 40.)

Hirotani, K., Pu, H.-Y., Lin, L. C.-C., et al., *Lepton Acceleration in the Vicinity of the Event Horizon: High-energy and Very-high-energy Emissions from Rotating Black Holes with Various Masses*, ApJ, 833, 142 (2016). (Cited on page 38.)

Ho, L. C., *The Spectral Energy Distributions of Low-Luminosity Active Galactic Nuclei*, ApJ, 516, 672 (1999). (Cited on page 25.)

Hockney, R. W. & Eastwood, J. W., *Computer Simulation Using Particles*, 1st edn., Routledge (1988). (Cited on pages 121, 122, and 130.)

Hornstein, S. D., Matthews, K., Ghez, A. M., et al., *A Constant Spectral Index for Sagittarius A* during Infrared/X-Ray Intensity Variations*, ApJ, 667, 900 (2007). (Cited on page 31.)

Howes, G. G., Cowley, S. C., Dorland, W., et al., *Astrophysical Gyrokinetics: Basic Equations and Linear Theory*, ApJ, 651, 590 (2006). (Cited on page 121.)

Hughes, D. H., Robson, E. I., Dunlop, J. S., & Gear, W. K., *Thermal dust emission from quasars. I. Submillimetre spectral indices of radio-quiet quasars.*, MNRAS, 263, 607 (1993). (Cited on page 18.)

Hughes, S. A., Keeton, Charles R., I., Walker, P., et al., *Finding black holes in numerical space-times*, Phys. Rev. D, 49, 4004 (1994). (Cited on page 134.)

Huppenkothen, D., Bachetti, M., Stevens, A. L., et al., *Stingray: A Modern Python Library for Spectral Timing*, ApJ, 881, 39 (2019). (Cited on page 196.)

Ipser, J. R., *Electromagnetic Test Fields Around a Kerr-Metric Black Hole*, Phys. Rev. Lett., 27, 529 (1971). (Cited on page 78.)

Israel, W., *Event Horizons in Static Vacuum Space-Times*, Physical Review, 164, 1776 (1967). (Cited on page 51.)

Jackson, J. D., *Classical Electrodynamics*, 3rd edn., John Wiley & Sons (1998). (Cited on pages 114 and 117.)

Jacquemin-Ide, J., Lesur, G., & Ferreira, J., *Magnetic outflows from turbulent accretion disks. I. Vertical structure and secular evolution*, A&A, 647, A192 (2021). (Cited on page 178.)

Jansky, K. G., *Radio Waves from Outside the Solar System*, Nature, 132, 66 (1933). (Cited on page 11.)

Jennison, R. C. & Das Gupta, M. K., *Fine Structure of the Extra-terrestrial Radio Source Cygnus I*, Nature, 172, 996 (1953). (Cited on page 11.)

Ji, H., Yamada, M., Hsu, S., et al., *Magnetic reconnection with Sweet-Parker characteristics in two-dimensional laboratory plasmas*, Physics of Plasmas, 6, 1743 (1999). (Cited on page 183.)

Johannsen, T. & Psaltis, D., *Testing the No-hair Theorem with Observations in the Electromagnetic Spectrum. II. Black Hole Images*, ApJ, 718, 446 (2010). (Cited on page 198.)

Johnson, M. D., Fish, V. L., Doeleman, S. S., et al., *Resolved magnetic-field structure and variability near the event horizon of Sagittarius A**, Science, 350, 1242 (2015). (Cited on page 31.)

Jones, F. C., *Calculated Spectrum of Inverse-Compton-Scattered Photons*, Physical Review, 167, 1159 (1968). (Cited on pages 105 and 106.)

Kadler, M., Eisenacher, D., Ros, E., et al., *The blazar-like radio structure of the TeV source IC 310*, A&A, 538, L1 (2012). (Cited on page 34.)

Kagan, D., Milosavljević, M., & Spitkovsky, A., *A Flux Rope Network and Particle Acceleration in Three-dimensional Relativistic Magnetic Reconnection*, ApJ, 774, 41 (2013). (Cited on page 182.)

Kagan, D., Sironi, L., Cerutti, B., & Giannios, D., *Relativistic Magnetic Reconnection in Pair Plasmas and Its Astrophysical Applications*, Space Sci. Rev., 191, 545 (2015). (Cited on page 182.)

Kalos, M. H. & Whitlock, P. A., *Monte Carlo Methods*, 2nd edn., Wiley VCH (2008). (Cited on page 151.)

Katsoulakos, G. & Rieger, F. M., *Magnetospheric Gamma-Ray Emission in Active Galactic Nuclei*, ApJ, 852, 112 (2018). (Cited on page 39.)

Katsoulakos, G. & Rieger, F. M., *Gap-type Particle Acceleration in the Magnetospheres of Rotating Supermassive Black Holes*, ApJ, 895, 99 (2020). (Cited on pages 39 and 40.)

Katz, M. L., Kelley, L. Z., Dosopoulou, F., et al., *Probing massive black hole binary populations with LISA*, MNRAS, 491, 2301 (2020). (Cited on page 17.)

Kellermann, K. I., *Grote Reber's Observations on Cosmic Static*, ApJ, 525C, 371 (1999). (Cited on page 11.)

Kelner, S. R., Prosekin, A. Y., & Aharonian, F. A., *Synchro-Curvature Radiation of Charged Particles in the Strong Curved Magnetic Fields*, AJ, 149, 33 (2015). (Cited on page 115.)

Kerr, R. P., *Gravitational Field of a Spinning Mass as an Example of Algebraically Special Metrics*, Phys. Rev. Lett., 11, 237 (1963). (Cited on page 6.)

Kim, J. Y., Krichbaum, T. P., Lu, R. S., et al., *The limb-brightened jet of M87 down to the 7 Schwarzschild radii scale*, A&A, 616, A188 (2018). (Cited on pages 21 and 27.)

King, A. R., Lasota, J. P., & Kundt, W., *Black holes and magnetic fields*, Phys. Rev. D, 12, 3037 (1975). (Cited on page 81.)

Kirk, J. G., Lyubarsky, Y., & Petri, J. (2009), in *Neutron Stars and Pulsars*, ed., W. Becker, Vol. 357, Springer, 421. (Cited on page 21.)

Kisaka, S., Levinson, A., & Toma, K., *Comprehensive Analysis of Magnetospheric Gaps around Kerr Black Holes Using 1D GRPIC Simulations*, ApJ, 902, 80 (2020). (Cited on pages 43, 167, and 197.)

Koide, S., *Magnetic extraction of black hole rotational energy: Method and results of general relativistic magnetohydrodynamic simulations in Kerr space-time*, Phys. Rev. D, 67, 104010 (2003). (Cited on page 41.)

Komissarov, S. S., *Direct numerical simulations of the Blandford-Znajek effect*, MNRAS, 326, L41 (2001). (Cited on page 41.)

Komissarov, S. S., *Limit shocks of relativistic magnetohydrodynamics, Punsly's waveguide and the Blandford-Znajek solution*, arXiv e-prints, arXiv:0206076 (2002a). (Cited on page 98.)

Komissarov, S. S., *Time-dependent, force-free, degenerate electrodynamics*, MNRAS, 336, 759 (2002b). (Cited on pages 84 and 97.)

Komissarov, S. S., *Electrodynamics of black hole magnetospheres*, MNRAS, 350, 427 (2004a). (Cited on pages 54, 58, 66, 86, 87, 88, 92, and 147.)

Komissarov, S. S., *General relativistic magnetohydrodynamic simulations of monopole magnetospheres of black holes*, MNRAS, 350, 1431 (2004b). (Cited on page 41.)

Komissarov, S. S., *Observations of the Blandford-Znajek process and the magnetohydrodynamic Penrose process in computer simulations of black hole magnetospheres*, MNRAS, 359, 801 (2005). (Cited on page 41.)

Komissarov, S. S., *Blandford-Znajek Mechanism versus Penrose Process*, Journal of Korean Physical Society, 54, 2503 (2009). (Cited on page 93.)

Komissarov, S. S. & McKinney, J. C., *The 'Meissner effect' and the Blandford-Znajek mechanism in conductive black hole magnetospheres*, MNRAS, 377, L49 (2007). (Cited on pages 41, 81, 94, 159, and 172.)

Kormendy, J. & Bender, R., *The Double Nucleus and Central Black Hole of M31*, ApJ, 522, 772 (1999). (Cited on page 15.)

Kormendy, J. & Richstone, D., *Inward Bound—The Search For Supermassive Black Holes In Galactic Nuclei*, ARA&A, 33, 581 (1995). (Cited on page 15.)

Kristian, J., *Quasars as Events in the Nuclei of Galaxies: the Evidence from Direct Photographs*, ApJ, 179, L61 (1973). (Cited on page 13.)

Kruskal, M. D., *Maximal Extension of Schwarzschild Metric*, Physical Review, 119, 1743 (1960). (Cited on page 5.)

Kuchelmeister, D., Müller, T., Ament, M., Wunner, G., & Weiskopf, D., *GPU-based four-dimensional general-relativistic ray tracing*, Computer Physics Communications, 183, 2282 (2012). (Cited on page 189.)

Kulier, A., Ostriker, J. P., Natarajan, P., Lackner, C. N., & Cen, R., *Understanding Black Hole Mass Assembly via Accretion and Mergers at Late Times in Cosmological Simulations*, ApJ, 799, 178 (2015). (Cited on page 15.)

Kunz, M. W., Schekochihin, A. A., & Stone, J. M., *Firehose and Mirror Instabilities in a Collisionless Shearing Plasma*, Phys. Rev. Lett., 112, 205003 (2014). (Cited on page 43.)

Kuo, C. Y., Asada, K., Rao, R., et al., *Measuring Mass Accretion Rate onto the Supermassive Black Hole in M87 Using Faraday Rotation Measure with the Submillimeter Array*, ApJ, 783, L33 (2014). (Cited on page 25.)

Lacy, J. H., Townes, C. H., Geballe, T. R., & Hollenbach, D. J., *Observations of the motion and distribution of the ionized gas in the central parsec of the Galaxy. II.*, ApJ, 241, 132 (1980). (Cited on page 29.)

Laing, R. A., *The sidedness of jets and depolarization in powerful extragalactic radio sources*, Nature, 331, 149 (1988). (Cited on page 19.)

Landau, L. D. & Lifshitz, E. M., *The Classical Theory of Fields*, 4th edn., Butterworth-Heinemann (1980). (Cited on pages 117 and 118.)

Lasota, J. P., Gourgoulhon, E., Abramowicz, M., Tchekhovskoy, A., & Narayan, R., *Extracting black-hole rotational energy: The generalized Penrose process*, Phys. Rev. D, 89, 024041 (2014). (Cited on page 99.)

Lesur, G., *Magnetohydrodynamics of protoplanetary discs*, arXiv e-prints, arXiv:2007.15967 (2020). (Cited on page 22.)

LeVeque, R. J., *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-dependent Problems*, 1st edn., Society for Industrial and Applied Mathematics (2007). (Cited on page 131.)

Levinson, A., *Particle Acceleration and Curvature TeV Emission by Rotating, Supermassive Black Holes*, Phys. Rev. Lett., 85, 912 (2000). (Cited on pages 37 and 38.)

Levinson, A. (2006), in *Trends in Black Hole Research*, ed., P. V. Kreitler, Nova Science Publishers, 119. (Cited on page 98.)

Levinson, A. & Cerutti, B., *Particle-in-cell simulations of pair discharges in a starved magnetosphere of a Kerr black hole*, A&A, 616, A184 (2018). (Cited on pages 43, 149, 171, and 197.)

Levinson, A., Melrose, D., Judge, A., & Luo, Q., *Large-Amplitude, Pair-creating Oscillations in Pulsar and Black Hole Magnetospheres*, ApJ, 631, 456 (2005). (Cited on pages 40 and 166.)

Levinson, A. & Rieger, F., *Variable TeV Emission as a Manifestation of Jet Formation in M87?*, ApJ, 730, 123 (2011). (Cited on pages 38, 39, 149, 161, and 167.)

Levinson, A. & Segev, N., *Existence of steady gap solutions in rotating black hole magnetospheres*, Phys. Rev. D, 96, 123006 (2017). (Cited on page 40.)

Liska, M., Tchekhovskoy, A., & Quataert, E., *Large-scale poloidal magnetic field dynamo leads to powerful jets in GRMHD simulations of black hole accretion with toroidal field*, MNRAS, 494, 3656 (2020). (Cited on page 22.)

Livio, M., Ogilvie, G. I., & Pringle, J. E., *Extracting Energy from Black Holes: The Relative Importance of the Blandford-Znajek Mechanism*, ApJ, 512, 100 (1999). (Cited on page 22.)

Longair, M. S., *High Energy Astrophysics*, 3rd edn., Cambridge University Press (2011). (Cited on pages 6, 15, 20, and 26.)

Loureiro, N. F., Cowley, S. C., Dorland, W. D., Haines, M. G., & Schekochihin, A. A., *X-Point Collapse and Saturation in the Nonlinear Tearing Mode Reconnection*, Phys. Rev. Lett., 95, 235003 (2005). (Cited on page 184.)

Loureiro, N. F., Schekochihin, A. A., & Cowley, S. C., *Instability of current sheets and formation of plasmoid chains*, Physics of Plasmas, 14, 100703 (2007). (Cited on page 184.)

Lubow, S. H., Papaloizou, J. C. B., & Pringle, J. E., *Magnetic field dragging in accretion discs*, MNRAS, 267, 235 (1994). (Cited on page 178.)

Luminet, J. P., *Image of a spherical black hole with thin accretion disk.*, A&A, 75, 228 (1979). (Cited on pages 28, 198, and 200.)

Lusso, E. & Risaliti, G., *The Tight Relation between X-Ray and Ultraviolet Luminosity of Quasars*, ApJ, 819, 154 (2016). (Cited on page 18.)

Lynden-Bell, D., *Galactic Nuclei as Collapsed Old Quasars*, Nature, 223, 690 (1969). (Cited on pages 8 and 14.)

Lyubarsky, Y., *Asymptotic Structure of Poynting-Dominated Jets*, ApJ, 698, 1570 (2009). (Cited on page 21.)

Lyubarsky, Y. E., *On the relativistic magnetic reconnection*, MNRAS, 358, 113 (2005). (Cited on page 183.)

Lyutikov, M. & Uzdensky, D., *Dynamics of Relativistic Reconnection*, ApJ, 589, 893 (2003). (Cited on page 182.)

MacDonald, D. & Thorne, K. S., *Black-hole electrodynamics - an absolute-space/universal-time formulation*, MNRAS, 198, 345 (1982). (Cited on pages 41, 60, 92, and 96.)

Madejski, G. G. & Sikora, M., *Gamma-Ray Observations of Active Galactic Nuclei*, ARA&A, 54, 725 (2016). (Cited on page 33.)

MAGIC Collaboration, Ansoldi, S., Antonelli, L. A., et al., *Gamma-ray flaring activity of NGC1275 in 2016-2017 measured by MAGIC*, A&A, 617, A91 (2018). (Cited on pages 34 and 35.)

Mahlmann, J. F., Levinson, A., & Aloy, M. A., *Striped Blandford/Znajek jets from advection of small-scale magnetic field*, MNRAS, 494, 4203 (2020). (Cited on page 172.)

Makwana, K. D., Keppens, R., & Lapenta, G., *Two-way coupling of magnetohydrodynamic simulations with embedded particle-in-cell simulations*, Computer Physics Communications, 221, 81 (2017). (Cited on page 213.)

Malkan, M. A. & Sargent, W. L. W., *The ultraviolet excess of Seyfert 1 galaxies and quasars.*, ApJ, 254, 22 (1982). (Cited on page 14.)

Markarian, B. E., *Galaxies with an ultraviolet continuum.*, Astrofizika, 3, 55 (1967). (Cited on page 12.)

Marrone, D. P., Moran, J. M., Zhao, J.-H., & Rao, R., *An Unambiguous Detection of Faraday Rotation in Sagittarius A**, ApJ, 654, L57 (2007). (Cited on page 30.)

Martel, A. R., Ford, H. C., Tran, H. D., et al., *Coronagraphic Imaging of 3C 273 with the Advanced Camera for Surveys*, AJ, 125, 2964 (2003). (Cited on page 13.)

Matthews, T. A. & Sandage, A. R., *Optical Identification of 3C 48, 3C 196, and 3C 286 with Stellar Objects.*, ApJ, 138, 30 (1963). (Cited on page 12.)

May, M. M. & White, R. H., *Hydrodynamic Calculations of General-Relativistic Collapse*, Physical Review, 141, 1232 (1966). (Cited on page 5.)

McClintock, J. E., Narayan, R., Garcia, M. R., et al., *Multiwavelength Spectrum of the Black Hole XTE J1118+480 in Quiescence*, ApJ, 593, 435 (2003). (Cited on page 10.)

McClintock, J. E., Narayan, R., & Rybicki, G. B., *On the Lack of Thermal Emission from the Quiescent Black Hole XTE J1118+480: Evidence for the Event Horizon*, ApJ, 615, 402 (2004). (Cited on page 10.)

McKinney, J. C., *General relativistic magnetohydrodynamic simulations of the jet formation and large-scale propagation from black hole accretion systems*, MNRAS, 368, 1561 (2006). (Cited on page 42.)

McKinney, J. C. & Gammie, C. F., *A Measurement of the Electromagnetic Luminosity of a Kerr Black Hole*, ApJ, 611, 977 (2004). (Cited on pages 42, 43, 57, and 94.)

McKinney, J. C., Tchekhovskoy, A., & Blandford, R. D., *General relativistic magnetohydrodynamic simulations of magnetically choked accretion flows around black holes*, MNRAS, 423, 3083 (2012). (Cited on pages 42 and 172.)

Mehlhaff, J. M., Werner, G. R., Uzdensky, D. A., & Begelman, M. C., *Kinetic beaming in radiative relativistic magnetic reconnection: a mechanism for rapid gamma-ray flares in jets*, MNRAS, 498, 799 (2020). (Cited on page 193.)

Menon, G. & Dermer, C. D., *A class of exact solutions to the force-free, axisymetric, stationary magnetosphere of a Kerr black hole*, General Relativity and Gravitation, 39, 785 (2007). (Cited on page 94.)

Merloni, A., Heinz, S., & di Matteo, T., *A Fundamental Plane of black hole activity*, MNRAS, 345, 1057 (2003). (Cited on page 19.)

Mertens, F., Lobanov, A. P., Walker, R. C., & Hardee, P. E., *Kinematics of the jet in M 87 on scales of 100-1000 Schwarzschild radii*, A&A, 595, A54 (2016). (Cited on pages 21 and 27.)

Michel, F. C., *Rotating Magnetosphere: a Simple Relativistic Model*, ApJ, 180, 207 (1973a). (Cited on page 90.)

Michel, F. C., *Rotating Magnetospheres: an Exact 3-D Solution*, ApJ, 180, L133 (1973b). (Cited on page 90.)

Michel, F. C. & Li, H., *Electrodynamics of neutron stars*, Phys. Rep., 318, 227 (1999). (Cited on page 95.)

Michell, J., *On the Means of Discovering the Distance, Magnitude, &c. of the Fixed Stars, in Consequence of the Diminution of the Velocity of Their Light, in Case Such a Diminution Should be Found to Take Place in any of Them, and Such Other Data Should be Procured from Observations, as Would be Farther Necessary for That Purpose. By the Rev. John Michell, B. D. F. R. S. In a Letter to Henry Cavendish, Esq. F. R. S. and A. S.*, Philosophical Transactions of the Royal Society of London Series I, 74, 35 (1784). (Cited on page 3.)

Mignone, A., Bodo, G., Vaidya, B., & Mattia, G., *A Particle Module for the PLUTO Code. I. An Implementation of the MHD-PIC Equations*, ApJ, 859, 13 (2018). (Cited on page 122.)

Miller-Jones, J. C. A., Bahramian, A., Orosz, J. A., et al., *Cygnus X-1 contains a 21–solar mass black hole—Implications for massive star winds*, Science, 371, 1046 (2021). (Cited on page 10.)

Misner, C. W., *Interpretation of Gravitational-Wave Observations*, Phys. Rev. Lett., 28, 994 (1972). (Cited on page 57.)

Miyoshi, M., Moran, J., Herrnstein, J., et al., *Evidence for a black hole from high rotation velocities in a sub-parsec region of NGC4258*, Nature, 373, 127 (1995). (Cited on pages 16 and 17.)

Moran, J. M., Greenhill, L. J., & Herrnstein, J. R., *Observational Evidence for Massive Black Holes in the Centers of Active Galaxies*, Journal of Astrophysics and Astronomy, 20, 165 (1999). (Cited on page 16.)

Mościbrodzka, M., Gammie, C. F., Dolence, J. C., & Shiokawa, H., *Pair Production in Low-luminosity Galactic Nuclei*, ApJ, 735, 9 (2011). (Cited on pages 43, 149, and 161.)

Narayan, R. & Heyl, J. S., *On the Lack of Type I X-Ray Bursts in Black Hole X-Ray Binaries: Evidence for the Event Horizon?*, ApJ, 574, L139 (2002). (Cited on page 10.)

Narayan, R., Igumenshchev, I. V., & Abramowicz, M. A., *Magnetically Arrested Disk: an Energetically Efficient Accretion Flow*, PASJ, 55, L69 (2003). (Cited on pages 42 and 172.)

Narayan, R., Johnson, M. D., & Gammie, C. F., *The Shadow of a Spherically Accreting Black Hole*, ApJ, 885, L33 (2019). (Cited on page 198.)

Narayan, R. & Yi, I., *Advection-dominated Accretion: A Self-similar Solution*, ApJ, 428, L13 (1994). (Cited on page 24.)

Narayan, R. & Yi, I., *Advection-dominated Accretion: Underfed Black Holes and Neutron Stars*, ApJ, 452, 710 (1995). (Cited on page 24.)

Neilsen, J., Nowak, M. A., Gammie, C., et al., *A Chandra/HETGS Census of X-Ray Variability from Sgr A* during 2012*, ApJ, 774, 42 (2013). (Cited on page 31.)

Neronov, A. & Aharonian, F. A., *Production of TeV Gamma Radiation in the Vicinity of the Supermassive Black Hole in the Giant Radio Galaxy M87*, ApJ, 671, 85 (2007). (Cited on pages 38, 39, 94, and 161.)

Novikov, I. D. & Thorne, K. S. (1973), in *Black Holes (Les Astres Occlus)*, ed., C. DeWitt & B. S. DeWitt, New York : Gordon and Breach, 343–450. (Cited on page 8.)

Numata, R., Howes, G. G., Tatsuno, T., Barnes, M., & Dorland, W., *AstroGK: Astrophysical gyrokinetics code*, Journal of Computational Physics, 229, 9347 (2010). (Cited on page 122.)

Oppenheimer, J. R. & Snyder, H., *On Continued Gravitational Contraction*, Physical Review, 56, 455 (1939). (Cited on page 5.)

Oppenheimer, J. R. & Volkoff, G. M., *On Massive Neutron Cores*, Physical Review, 55, 374 (1939). (Cited on page 4.)

Orosz, J. A., McClintock, J. E., Aufdenberg, J. P., et al., *The Mass of the Black Hole in Cygnus X-1*, ApJ, 742, 84 (2011). (Cited on page 10.)

Owen, F. N., Eilek, J. A., & Kassim, N. E., *M87 at 90 Centimeters: A Different Picture*, ApJ, 543, 611 (2000). (Cited on page 24.)

Padovani, P., Alexander, D. M., Assef, R. J., et al., *Active galactic nuclei: what's in a name?*, A&A Rev., 25, 2 (2017). (Cited on pages 18 and 33.)

Palmroth, M., Ganse, U., Pfau-Kempf, Y., et al., *Vlasov methods in space physics and astrophysics*, Living Reviews in Computational Astrophysics, 4, 1 (2018). (Cited on page 121.)

Pan, Z. & Yu, C., *Fourth-order split monopole perturbation solutions to the Blandford-Znajek mechanism*, Phys. Rev. D, 91, 064067 (2015). (Cited on page 94.)

Parfrey, K., Giannios, D., & Beloborodov, A. M., *Black hole jets without large-scale net magnetic flux.*, MNRAS, 446, L61 (2015). (Cited on pages 94 and 172.)

Parfrey, K., Philippov, A., & Cerutti, B., *First-Principles Plasma Simulations of Black-Hole Jet Launching*, Phys. Rev. Lett., 122, 035101 (2019). (Cited on pages 44, 134, 140, 159, 165, 166, 170, and 176.)

Parker, E. N., *Sweet's Mechanism for Merging Magnetic Fields in Conducting Fluids*, J. Geophys. Res., 62, 509 (1957). (Cited on page 183.)

Pearson, T. J., Unwin, S. C., Cohen, M. H., et al., *Superluminal expansion of quasar 3C273*, Nature, 290, 365 (1981). (Cited on page 19.)

Penrose, R., *Gravitational Collapse and Space-Time Singularities*, Phys. Rev. Lett., 14, 57 (1965). (Cited on page 6.)

Penrose, R., *"Golden Oldie": Gravitational Collapse: The Role of General Relativity*, General Relativity and Gravitation, 7, 1141 (2002). (Cited on page 53.)

Penrose, R. & Floyd, R. M., *Extraction of Rotational Energy from a Black Hole*, Nature Physical Science, 229, 177 (1971). (Cited on page 56.)

Peterson, B. M. & Bentz, M. C., *Black hole masses from reverberation mapping*, New A Rev., 50, 796 (2006). (Cited on page 16.)

Philippov, A., Timokhin, A., & Spitkovsky, A., *Origin of Pulsar Radio Emission*, Phys. Rev. Lett., 124, 245101 (2020). (Cited on page 166.)

Podurets, M. A., *The Collapse of a Star with Back Pressure Taken into Account*, Soviet Physics Doklady, 9, 1 (1964). (Cited on page 5.)

Ponti, G., George, E., Scaringi, S., et al., *A powerful flare from Sgr A* confirms the synchrotron nature of the X-ray emission*, MNRAS, 468, 2447 (2017). (Cited on page 31.)

Porth, O., Chatterjee, K., Narayan, R., et al., *The Event Horizon General Relativistic Magnetohydrodynamic Code Comparison Project*, ApJS, 243, 26 (2019). (Cited on page 43.)

Porth, O., Mizuno, Y., Younsi, Z., & Fromm, C. M., *Flares in the Galactic Centre - I. Orbiting flux tubes in magnetically arrested black hole accretion discs*, MNRAS, 502, 2023 (2021). (Cited on page 32.)

Pottschmidt, K. & Konig, M. (1997), in *Astronomical Time Series*, ed., D. Maoz, A. Sternberg, & E. M. Leibowitz, Vol. 218, Springer, 187. (Cited on pages 8 and 10.)

Press, W. H. & Teukolsky, S. A., *Floating Orbits, Superradiant Scattering and the Black-hole Bomb*, Nature, 238, 211 (1972). (Cited on page 57.)

Press, W. H. & Teukolsky, S. A., *Perturbations of a Rotating Black Hole. II. Dynamical Stability of the Kerr Metric*, ApJ, 185, 649 (1973). (Cited on page 6.)

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P., *Numerical recipes in FORTRAN. The art of scientific computing*, 2nd edn., Cambridge University Press (1992). (Cited on page 131.)

Price, R. H., *Nonspherical Perturbations of Relativistic Gravitational Collapse. I. Scalar and Gravitational Perturbations*, Phys. Rev. D, 5, 2419 (1972a). (Cited on page 6.)

Price, R. H., *Nonspherical Perturbations of Relativistic Gravitational Collapse. II. Integer-Spin, Zero-Rest-Mass Fields*, Phys. Rev. D, 5, 2439 (1972b). (Cited on page 6.)

Prieto, M. A., Fernández-Ontiveros, J. A., Markoff, S., Espada, D., & González-Martín, O., *The central parsecs of M87: jet emission and an elusive accretion disc*, MNRAS, 457, 3801 (2016). (Cited on pages 18, 24, 25, and 161.)

Pritchett, P. (2003), in *Space Plasma Simulation*, ed., J. Büchner et al., Springer, 1–24. (Cited on pages 122 and 131.)

Ptitsyna, K. & Neronov, A., *Particle acceleration in the vacuum gaps in black hole magnetospheres*, A&A, 593, A8 (2016). (Cited on page 39.)

Punsly, B., *Black hole gravitohydromagnetics*, 1st edn., Springer (2001). (Cited on pages 41 and 98.)

Punsly, B. & Coroniti, F. V., *Ergosphere-driven Winds*, ApJ, 354, 583 (1990a). (Cited on page 41.)

Punsly, B. & Coroniti, F. V., *Relativistic Winds from Pulsar and Black Hole Magnetospheres*, ApJ, 350, 518 (1990b). (Cited on page 41.)

Qin, H., Zhang, S., Xiao, J., et al., *Why is Boris algorithm so good?*, Physics of Plasmas, 20, 084503 (2013). (Cited on page 126.)

Quataert, E., *Radiatively Inefficient Accretion Flow Models of Sgr A**, Astronomische Nachrichten Supplement, 324, 435 (2003). (Cited on page 25.)

Quataert, E. & Gruzinov, A., *Constraining the Accretion Rate onto Sagittarius A* Using Linear Polarization*, ApJ, 545, 842 (2000a). (Cited on page 25.)

Quataert, E. & Gruzinov, A., *Convection-dominated Accretion Flows*, ApJ, 539, 809 (2000b). (Cited on page 24.)

Rawlings, S. & Saunders, R., *Evidence for a common central-engine mechanism in all extragalactic radio sources*, Nature, 349, 138 (1991). (Cited on pages 18, 19, and 22.)

Reber, G., *Cosmic Static*, ApJ, 100, 279 (1944). (Cited on page 11.)

Rees, M. J., *Appearance of Relativistically Expanding Radio Sources*, Nature, 211, 468 (1966). (Cited on page 19.)

Rees, M. J., *Black Hole Models for Active Galactic Nuclei*, ARA&A, 22, 471 (1984). (Cited on page 14.)

Reid, M. J. & Brunthaler, A., *The Proper Motion of Sagittarius A*. II. The Mass of Sagittarius A**, ApJ, 616, 872 (2004). (Cited on page 29.)

Reid, M. J. & Brunthaler, A., *The Proper Motion of Sagittarius A*. III. The Case for a Supermassive Black Hole*, ApJ, 892, 39 (2020). (Cited on page 29.)

Reid, M. J., Readhead, A. C. S., Vermeulen, R. C., & Treuhaft, R. N., *The Proper Motion of Sagittarius A*. I. First VLBA Results*, ApJ, 524, 816 (1999). (Cited on page 29.)

Ressler, S. M., Quataert, E., White, C. J., & Blaes, O., *Magnetically modified spherical accretion in GRMHD: reconnection-driven convection and jet propagation*, MNRAS, 504, 6076 (2021). (Cited on page 42.)

Ressler, S. M., White, C. J., Quataert, E., & Stone, J. M., *Ab Initio Horizon-scale Simulations of Magnetically Arrested Accretion in Sagittarius A* Fed by Stellar Winds*, ApJ, 896, L6 (2020). (Cited on page 42.)

Reynolds, C. S., *The spin of supermassive black holes*, Classical and Quantum Gravity, 30, 244004 (2013). (Cited on page 49.)

Reynolds, C. S., *Observational Constraints on Black Hole Spin*, arXiv e-prints, arXiv:2011.08948 (2020). (Cited on pages 16 and 49.)

Reynolds, C. S., Di Matteo, T., Fabian, A. C., Hwang, U., & Canizares, C. R., *The 'quiescent' black hole in M87*, MNRAS, 283, L111 (1996). (Cited on pages 18, 23, and 24.)

Rhoades, C. E. & Ruffini, R., *Maximum Mass of a Neutron Star*, Phys. Rev. Lett., 32, 324 (1974). (Cited on page 5.)

Rieger, F. & Levinson, A., *Radio Galaxies at VHE Energies*, Galaxies, 6, 116 (2018). (Cited on pages 34, 37, and 39.)

Rieger, F. M. & Aharonian, F., *Probing the Central Black Hole in M87 with Gamma-Rays*, Modern Physics Letters A, 27, 1230030 (2012). (Cited on pages 37 and 39.)

Ripperda, B., Bacchini, F., & Philippov, A. A., *Magnetic Reconnection and Hot Spot Formation in Black Hole Accretion Disks*, ApJ, 900, 100 (2020). (Cited on pages 32, 43, 172, and 196.)

Ripperda, B., Bacchini, F., Teunissen, J., et al., *A Comprehensive Comparison of Relativistic Particle Integrators*, ApJS, 235, 21 (2018). (Cited on page 126.)

Robinson, D. C., *Uniqueness of the Kerr Black Hole*, Phys. Rev. Lett., 34, 905 (1975). (Cited on page 51.)

Ruderman, M. A. & Sutherland, P. G., *Theory of pulsars: polar gaps, sparks, and coherent microwave radiation.*, ApJ, 196, 51 (1975). (Cited on page 37.)

Runge, J. & Walker, S. A., *Probing within the Bondi radius of the ultramassive black hole in NGC 1600*, MNRAS, 502, 5487 (2021). (Cited on page 25.)

Ryan, B. R., Ressler, S. M., Dolence, J. C., Gammie, C., & Quataert, E., *Two-temperature GRRMHD Simulations of M87*, ApJ, 864, 126 (2018). (Cited on page 43.)

Rybicki, G. B. & Lightman, A. P., *Radiative Processes in Astrophysics*, 1st edn., John Wiley & Sons (1986). (Cited on pages 26, 103, and 113.)

Sahakyan, N., Baghmanyan, V., & Zargaryan, D., *Fermi-LAT observation of nonblazar AGNs*, A&A, 614, A6 (2018). (Cited on page 34.)

Salpeter, E. E., *Accretion of Interstellar Matter by Massive Objects.*, ApJ, 140, 796 (1964). (Cited on page 6.)

Samtaney, R., Loureiro, N. F., Uzdensky, D. A., Schekochihin, A. A., & Cowley, S. C., *Formation of Plasmoid Chains in Magnetic Reconnection*, Phys. Rev. Lett., 103, 105004 (2009). (Cited on page 184.)

Sandage, A., *The Existence of a Major New Constituent of the Universe: the Quasistellar Galaxies.*, ApJ, 141, 1560 (1965). (Cited on page 12.)

Sargent, W. L. W., Young, P. J., Boksenberg, A., et al., *Dynamical evidence for a central mass concentration in the galaxy M87.*, ApJ, 221, 731 (1978). (Cited on page 15.)

Scharlemann, E. T. & Wagoner, R. V., *Aligned Rotating Magnetospheres. General Analysis*, ApJ, 182, 951 (1973). (Cited on page 90.)

Schekochihin, A. A., *MHD Turbulence: A Biased Review*, arXiv e-prints, arXiv:2010.00699 (2020). (Cited on page 184.)

Schmidt, M., *3C 273 : A Star-Like Object with Large Red-Shift*, Nature, 197, 1040 (1963). (Cited on page 12.)

Schödel, R., Ott, T., Genzel, R., et al., *A star in a 15.2-year orbit around the supermassive black hole at the centre of the Milky Way*, Nature, 419, 694 (2002). (Cited on page 30.)

Schwarzschild, K., *Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie*, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin) (1916). (Cited on page 3.)

Seyfert, C. K., *Nuclear Emission in Spiral Nebulae.*, ApJ, 97, 28 (1943). (Cited on page 12.)

Shakura, N. I. & Sunyaev, R. A., *Reprint of 1973A&A....24..337S. Black holes in binary systems. Observational appearance.*, A&A, 500, 33 (1973). (Cited on pages 8 and 25.)

Shapiro, S. L. & Teukolsky, A. A., *Black Holes, White Dwarfs and Neutron Stars: The Physics of Compact Objects*, 1st edn., John Wiley & Sons (1983). (Cited on pages 5 and 8.)

Shields, G. A., *A Brief History of Active Galactic Nuclei*, PASP, 111, 661 (1999). (Cited on page 10.)

Sikora, M., Begelman, M. C., & Rees, M. J., *Comptonization of Diffuse Ambient Radiation by a Relativistic Jet: The Source of Gamma Rays from Blazars?*, ApJ, 421, 153 (1994). (Cited on page 33.)

Sikora, M. & Madejski, G., *On Pair Content and Variability of Subparsec Jets in Quasars*, ApJ, 534, 109 (2000). (Cited on page 23.)

Sironi, L. & Cerutti, B. (2017), in *Modelling Pulsar Wind Nebulae*, ed., D. F. Torres, Vol. 446, Springer, 1–27. (Cited on pages 122, 123, 128, and 131.)

Sironi, L., Giannios, D., & Petropoulou, M., *Plasmoids in relativistic reconnection, from birth to adulthood: first they grow, then they go*, MNRAS, 462, 48 (2016). (Cited on page 184.)

Sironi, L. & Spitkovsky, A., *Relativistic Reconnection: An Efficient Source of Non-thermal Particles*, ApJ, 783, L21 (2014). (Cited on page 182.)

Sądowski, A., Wielgus, M., Narayan, R., et al., *Radiative, two-temperature simulations of low-luminosity black hole accretion flows in general relativity*, MNRAS, 466, 705 (2017). (Cited on page 43.)

Sol, H., Pelletier, G., & Asseo, E., *Two-flow model for extragalactic radio jets.*, MNRAS, 237, 411 (1989). (Cited on page 23.)

Soltan, A., *Masses of quasars.*, MNRAS, 200, 115 (1982). (Cited on page 15.)

Sowers, J. W., Gies, D. R., Bagnuolo, W. G., et al., *Tomographic Analysis of Hα Profiles in HDE 226868/Cygnus X-1*, ApJ, 506, 424 (1998). (Cited on pages 9 and 10.)

Spitkovsky, A., *Time-dependent Force-free Pulsar Magnetospheres: Axisymmetric and Oblique Rotators*, ApJ, 648, L51 (2006). (Cited on page 177.)

Springel, V., *High Performance Computing and Numerical Modelling*, Saas-Fee Advanced Course, 43, 251 (2016). (Cited on pages 124 and 136.)

Stone, J. M. & Pringle, J. E., *Magnetohydrodynamical non-radiative accretion flows in two dimensions*, MNRAS, 322, 461 (2001). (Cited on page 25.)

Stone, J. M., Pringle, J. E., & Begelman, M. C., *Hydrodynamical non-radiative accretion flows in two dimensions*, MNRAS, 310, 1002 (1999). (Cited on page 25.)

Strang, G., *On the Construction and Comparison of Difference Schemes*, SIAM Journal on Numerical Analysis, 5, 506 (1968). (Cited on page 136.)

Sturrock, P. A., *A Model of Pulsars*, ApJ, 164, 529 (1971). (Cited on page 37.)

Svensson, R., *Electron-Positron Pair Equilibria in Relativistic Plasmas*, ApJ, 258, 335 (1982). (Cited on page 113.)

Sweet, P. A. (1958), in *Electromagnetic Phenomena in Cosmical Physics*, ed., B. Lehnert, Vol. 6, Cambridge University Press, 123. (Cited on page 183.)

Swope, W. C., Andersen, H. C., Berens, P. H., & Wilson, K. R., *A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters*, J. Chem. Phys., 76, 637 (1982). (Cited on page 123.)

Takahashi, M., Nitta, S., Tatematsu, Y., & Tomimatsu, A., *Magnetohydrodynamic Flows in Kerr Geometry: Energy Extraction from Black Holes*, ApJ, 363, 206 (1990). (Cited on pages 40 and 98.)

Takahashi, R., *Horizon-penetrating transonic accretion discs around rotating black holes*, MNRAS, 382, 567 (2007). (Cited on pages 58 and 147.)

Tamburini, M., Pegoraro, F., Di Piazza, A., Keitel, C. H., & Macchi, A., *Radiation reaction effects on radiation pressure acceleration*, New Journal of Physics, 12, 123005 (2010). (Cited on page 126.)

Tanaka, Y., Nandra, K., Fabian, A. C., et al., *Gravitationally redshifted emission implying an accretion disk and massive black hole in the active galaxy MCG-6-30-15*, Nature, 375, 659 (1995). (Cited on page 16.)

Tchekhovskoy, A., Narayan, R., & McKinney, J. C., *Black Hole Spin and The Radio Loud/Quiet Dichotomy of Active Galactic Nuclei*, ApJ, 711, 50 (2010). (Cited on pages 94 and 173.)

Tchekhovskoy, A., Narayan, R., & McKinney, J. C., *Efficient generation of jets from magnetically arrested accretion on a rapidly spinning black hole*, MNRAS, 418, L79 (2011). (Cited on pages 22 and 42.)

Teo, E., *Spherical Photon Orbits Around a Kerr Black Hole*, General Relativity and Gravitation, 35, 1909 (2003). (Cited on page 205.)

Thoelecke, K., Takahashi, M., & Tsuruta, S., *Perturbing Monopolar Force-Free Magnetospheres to Slowly Rotating Black Holes*, arXiv e-prints, arXiv:1901.01245 (2019). (Cited on page 93.)

Thorne, K. S., *Black Holes & Time Warps – Einstein's Outrageous Legacy*, 1st edn., W. W. Norton & Company (1995). (Cited on pages 3, 4, and 5.)

Thorne, K. S. & MacDonald, D., *Electrodynamics in Curved Spacetime - 3+1 Formulation*, MNRAS, 198, 339 (1982). (Cited on pages 60 and 69.)

Thorne, K. S., Price, R. H., & MacDonald, D. A., *Black holes: The membrane paradigm*, 1st edn., Yale University Press (1986). (Cited on page 95.)

Timokhin, A. N., *Time-dependent pair cascades in magnetospheres of neutron stars - I. Dynamics of the polar cap cascade with no particle supply from the neutron star surface*, MNRAS, 408, 2092 (2010). (Cited on page 43.)

Timokhin, A. N. & Arons, J., *Current flow and pair creation at low altitude in rotation-powered pulsars' force-free magnetospheres: space charge limited flow*, MNRAS, 429, 20 (2013). (Cited on page 43.)

Timokhin, A. N. & Harding, A. K., *On the Maximum Pair Multiplicity of Pulsar Cascades*, ApJ, 871, 12 (2019). (Cited on page 162.)

Titarchuk, L. & Shaposhnikov, N., *How to Distinguish Neutron Star and Black Hole X-Ray Binaries? Spectral Index and Quasi-Periodic Oscillation Frequency Correlation*, ApJ, 626, 298 (2005). (Cited on page 9.)

Toma, K. & Takahara, F., *Electromotive force in the Blandford-Znajek process*, MNRAS, 442, 2855 (2014). (Cited on page 87.)

Trippe, S., Paumard, T., Ott, T., et al., *A polarized infrared flare from Sagittarius A* and the signatures of orbiting plasma hotspots*, MNRAS, 375, 764 (2007). (Cited on page 31.)

Trotter, H. F., *On the Product of Semi-Groups of Operators*, Proceedings of the American Mathematical Society, 10, 4. (Cited on page 136.)

Umeda, T., Togano, K., & Ogino, T., *Two-dimensional full-electromagnetic Vlasov code with conservative scheme and its application to magnetic reconnection*, Computer Physics Communications, 180, 365 (2009). (Cited on page 121.)

Urry, C. M. & Padovani, P., *Unified Schemes for Radio-Loud Active Galactic Nuclei*, PASP, 107, 803 (1995). (Cited on page 34.)

Uzdensky, D. A., *Force-Free Magnetosphere of an Accretion Disk-Black Hole System. I. Schwarzschild Geometry*, ApJ, 603, 652 (2004). (Cited on page 92.)

Uzdensky, D. A., *Force-Free Magnetosphere of an Accretion Disk-Black Hole System. II. Kerr Geometry*, ApJ, 620, 889 (2005). (Cited on page 92.)

Uzdensky, D. A. & Goodman, J., *Statistical Description of a Magnetized Corona above a Turbulent Accretion Disk*, ApJ, 682, 608 (2008). (Cited on page 212.)

Uzdensky, D. A. & Kulsrud, R. M., *Two-dimensional numerical simulation of the resistive reconnection layer*, Physics of Plasmas, 7, 4018 (2000). (Cited on page 183.)

Uzdensky, D. A., Loureiro, N. F., & Schekochihin, A. A., *Fast Magnetic Reconnection in the Plasmoid-Dominated Regime*, Phys. Rev. Lett., 105, 235002 (2010). (Cited on page 184.)

van Marle, A. J., Casse, F., & Marcowith, A., *On magnetic field amplification and particle acceleration near non-relativistic astrophysical shocks: particles in MHD cells simulations*, MNRAS, 473, 3394 (2018). (Cited on page 122.)

Verlet, L., *Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules*, Physical Review, 159, 98 (1967). (Cited on page 123.)

Villasenor, J. & Buneman, O., *Rigorous charge conservation for local electromagnetic field solvers*, Computer Physics Communications, 69, 306 (1992). (Cited on page 131.)

Vincent, F. H., Abramowicz, M. A., Zdziarski, A. A., et al., *Multi-wavelength torus-jet model for Sagittarius A**, A&A, 624, A52 (2019). (Cited on page 199.)

Vincent, F. H., Gourgoulhon, E., & Novak, J., *3+1 geodesic equation and images in numerical spacetimes*, Classical and Quantum Gravity, 29, 245005 (2012). (Cited on page 135.)

Vincent, F. H., Paumard, T., Gourgoulhon, E., & Perrin, G., *GYOTO: a new general relativistic ray-tracing code*, Classical and Quantum Gravity, 28, 225011 (2011). (Cited on page 189.)

Vincent, F. H., Wielgus, M., Abramowicz, M. A., et al., *Geometric modeling of M87* as a Kerr black hole or a non-Kerr compact object*, A&A, 646, A37 (2021). (Cited on pages 28 and 199.)

Vincent, S. & LeBohec, S., *Monte Carlo simulation of electromagnetic cascades in black hole magnetosphere*, MNRAS, 409, 1183 (2010). (Cited on page 39.)

Wald, R., *On uniqueness of the Kerr-Newman black holes.*, Journal of Mathematical Physics, 13, 490 (1972). (Cited on page 78.)

Wald, R., *Black hole in a uniform magnetic field*, Phys. Rev. D, 10, 1680 (1974). (Cited on page 80.)

Walker, R. C., Hardee, P. E., Davies, F. B., Ly, C., & Junor, W., *The Structure and Dynamics of the Subparsec Jet in M87 Based on 50 VLBA Observations over 17 Years at 43 GHz*, ApJ, 855, 128 (2018). (Cited on pages 21, 27, and 204.)

Walsh, J. L., Barth, A. J., Ho, L. C., & Sarzi, M., *The M87 Black Hole Mass from Gas-dynamical Models of Space Telescope Imaging Spectrograph Observations*, ApJ, 770, 86 (2013). (Cited on page 28.)

Wandel, A. & Mushotzky, R. F., *Observational Determination of the Masses of Active Galactic Nuclei*, ApJ, 306, L61 (1986). (Cited on page 14.)

Wang, C.-C. & Zhou, H.-Y., *Determination of the intrinsic velocity field in the M87 jet*, MNRAS, 395, 301 (2009). (Cited on page 35.)

Wang, J.-M., Li, Y.-R., Wang, J.-C., & Zhang, S., *Spins of the Supermassive Black Hole in M87: New Constraints from TeV Observations*, ApJ, 676, L109 (2008). (Cited on page 39.)

Wardle, J., *The Variable Rotation Measure Distribution in 3C 273 on Parsec Scales*, Galaxies, 6, 5 (2018). (Cited on page 19.)

Wardle, J. F. C., Homan, D. C., Ojha, R., & Roberts, D. H., *Electron-positron jets associated with the quasar 3C279*, Nature, 395, 457 (1998). (Cited on page 23.)

Webster, B. L. & Murdin, P., *Cygnus X-1-a Spectroscopic Binary with a Heavy Companion?*, Nature, 235, 37 (1972). (Cited on page 9.)

Werner, G. R., Uzdensky, D. A., Cerutti, B., Nalewajko, K., & Begelman, M. C., *The Extent of Power-law Energy Spectra in Collisionless Relativistic Magnetic Reconnection in Pair Plasmas*, ApJ, 816, L8 (2016). (Cited on page 182.)

Wheeler, J. A., *Superdense Stars*, ARA&A, 4, 393 (1966). (Cited on page 5.)

Whysong, D. & Antonucci, R., *Thermal Emission as a Test for Hidden Nuclei in Nearby Radio Galaxies*, ApJ, 602, 116 (2004). (Cited on page 39.)

Wilms, J., Reynolds, C. S., Begelman, M. C., et al., *XMM-EPIC observation of MCG-6-30-15: direct evidence for the extraction of energy from a spinning black hole?*, MNRAS, 328, L27 (2001). (Cited on page 23.)

Winske, D., Yin, L., Omidi, N., & et al. (2003), in *Space Plasma Simulation*, ed., J. Büchner, C. Dum, & M. Scholer, Vol. 615, Springer, 136–165. (Cited on page 122.)

Wong, G. N., Ryan, B. R., & Gammie, C. F., *Pair Drizzle around Sub-Eddington Supermassive Black Holes*, ApJ, 907, 73 (2021). (Cited on page 43.)

Wong, K.-W., Irwin, J. A., Yukita, M., et al., *Resolving the Bondi Accretion Flow toward the Supermassive Black Hole of NGC 3115 with Chandra*, ApJ, 736, L23 (2011). (Cited on page 25.)

Yamada, M., Kulsrud, R., & Ji, H., *Magnetic reconnection*, Reviews of Modern Physics, 82, 603 (2010). (Cited on page 182.)

Yee, K., *Numerical solution of inital boundary value problems involving Maxwell's equations in isotropic media*, IEEE Transactions on Antennas and Propagation, 14, 302 (1966). (Cited on page 128.)

Young, P. A. (2013), *The leapfrog method and other "symplectic" algorithms for integrating Newton's laws of motion*, URL: https://young.physics.ucsc.edu/115/leapfrog.pdf. Last visited on 2021/03/08. (Cited on page 124.)

Young, P. J., Westphal, J. A., Kristian, J., Wilson, C. P., & Landauer, F. P., *Evidence for a supermassive object in the nucleus of the galaxy M87 from SIT and CCD area photometry.*, ApJ, 221, 721 (1978). (Cited on page 15.)

Yuan, F. & Narayan, R., *Hot Accretion Flows Around Black Holes*, ARA&A, 52, 529 (2014). (Cited on page 24.)

Zdziarski, A. A., *Saturated Pair-Photon Cascades on Isotropic Background Photons*, ApJ, 335, 786 (1988). (Cited on page 112.)

Zel'dovich, Y. B., *The Fate of a Star and the Evolution of Gravitational Energy Upon Accretion*, Soviet Physics Doklady, 9, 195 (1964). (Cited on page 6.)

Znajek, R. L., *Black hole electrodynamics and the Carter tetrad.*, MNRAS, 179, 457 (1977). (Cited on pages 92, 99, and 100.)

Zweibel, E. G. & Yamada, M., *Magnetic Reconnection in Astrophysical and Laboratory Plasmas*, ARA&A, 47, 291 (2009). (Cited on page 182.)

Tarte tatin (Conticini et al., 2012)

Ingredients (6 servings)

- 1 sheet all-butter puff pastry
- 6 apples
- 100 g golden caster sugar
- 85 g unsalted butter

Instructions

1. Roll the pastry to a 3 mm-thick round on a lightly floured surface and cut a 24 cm circle, using a plate as a guide.
2. Poke holes all over with a fork, place on a baking sheet, then set aside.
3. Heat oven to 180° C.
4. Peel, quarter and core the apples.
5. In a large saucepan, over medium heat, pour the sugar and 3 tablespoons of water evenly and cook until light amber in color, stirring to help melt any lumps, during about 5-7 minutes. Add the butter, stirring constantly until the color is a creamy light brown.
6. Add the apples, stirring until they are coated in a thick layer of caramel. Cook for about 15-20 minutes, turning the apples constantly so that they bathe in the caramel. Remove from the heat when the caramel has reduced and little remains in the bottom of the pan.
7. Arrange the apple quarters very tightly in a circle around the edge of the dish first, then fill in the middle in a similar fashion. Gently press with your hands to ensure there are no gaps. Pour the remaining caramel over the top.
8. Lay the circle of puff pastry on top. Tuck the puff pastry down the sides of the pan. Bake for 45-50 minutes, until the pastry is golden brown and firm.
9. Allow to cool to room temperature for 1 hour before running a knife around the edge of the dish and inverting it onto a large serving plate that is deep enough to contain the juices. Serve with crème fraîche or vanilla ice cream.